

GenCore Version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 03:19:51 : Search time 2725.73 Seconds  
(without alignments)  
42.367 Million cell updates/sec

Title: US-09-380-826A-4  
Perfect score: 7  
Sequence: 1 tcttga 7

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues  
Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Geneml:\*

1: gb\_ba:\*

2: gb\_htg:\*

3: gb\_in:\*

4: gb\_com:\*

5: gb\_ov:\*

6: gb\_ph:\*

7: gb\_pl:\*

8: gb\_pr:\*

9: gb\_ro:\*

10: gb\_sts:\*

11: gb\_sy:\*

12: gb\_un:\*

13: gb\_vl:\*

14: gb\_wa:\*

15: em\_fun:\*

16: em\_hum:\*

17: em\_in:\*

18: em\_ov:\*

19: em\_or:\*

20: em\_pat:\*

21: em\_ph:\*

22: em\_pl:\*

23: em\_ro:\*

24: em\_sy:\*

25: em\_un:\*

26: em\_vl:\*

27: em\_wa:\*

28: em\_hum:\*

29: em\_in:\*

30: em\_ov:\*

31: em\_or:\*

32: em\_pat:\*

33: em\_ph:\*

34: em\_pl:\*

35: em\_ro:\*

36: em\_sy:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	7	100.0	10	6	AR062865	AR062865 Sequence
C 2	7	100.0	12	6	AX073580	AX073580 Sequence
C 3	7	100.0	13	6	A00432	A00432 Nucleotide
C 4	7	100.0	14	6	A25812	A25812 polynucleot
C 5	7	100.0	14	6	A40518	A40518 Sequence 55
C 6	7	100.0	14	6	A88271	A88271 Sequence 41
C 7	7	100.0	14	6	A88273	A88273 Sequence 42
C 8	7	100.0	14	6	A89045	A89045 Sequence 11
C 9	7	100.0	14	6	A90238	A90238 Sequence 41
C 10	7	100.0	14	6	A90240	A90240 Sequence 42
C 11	7	100.0	14	6	AR119028	AR119028 Sequence
C 12	7	100.0	14	6	AX007733	AX007733 Sequence
C 13	7	100.0	14	6	AX007734	AX007734 Sequence
C 14	7	100.0	14	6	AX009010	AX009010 Sequence
C 15	7	100.0	14	6	I26234	I26234 Sequence 19
C 16	7	100.0	15	6	A88272	A88272 Sequence 42
C 17	7	100.0	15	6	A88466	A88466 Sequence 61
C 18	7	100.0	15	6	A90239	A90239 Sequence 42
C 19	7	100.0	15	6	A90433	A90433 Sequence 61
C 20	7	100.0	15	6	AR033435	AR033435 Sequence
C 21	7	100.0	15	6	AR033480	AR033480 Sequence
C 22	7	100.0	15	6	AR113257	AR113257 Sequence
C 23	7	100.0	15	6	AR113302	AR113302 Sequence
C 24	7	100.0	15	6	AR132941	AR132941 Sequence
C 25	7	100.0	15	6	AR132942	AR132942 Sequence
C 26	7	100.0	15	6	AR133304	AR133304 Sequence
C 27	7	100.0	15	6	AR133305	AR133305 Sequence
C 28	7	100.0	15	6	AR133306	AR133306 Sequence
C 29	7	100.0	15	6	AR133307	AR133307 Sequence
C 30	7	100.0	15	6	AR133343	AR133343 Sequence
C 31	7	100.0	15	6	AR133344	AR133344 Sequence
C 32	7	100.0	15	6	AR133345	AR133345 Sequence
C 33	7	100.0	15	6	AR133714	AR133714 Sequence
C 34	7	100.0	15	6	AX007732	AX007732 Sequence
C 35	7	100.0	15	6	AX081111	AX081111 Sequence
C 36	7	100.0	15	6	I57664	I57664 Sequence 20
C 37	7	100.0	15	6	I57709	I57709 Sequence 24
C 38	7	100.0	16	6	A96937	A96937 Sequence 15
C 39	7	100.0	16	6	A97828	A97828 Sequence 10
C 40	7	100.0	16	6	AR084439	AR084439 Sequence
C 41	7	100.0	16	6	AR093881	AR093881 Sequence
C 42	7	100.0	16	6	AX133195	AX133195 Sequence
C 43	7	100.0	16	6	I34946	I34946 Sequence 32
C 44	7	100.0	16	6	I52073	I52073 Sequence 15
C 45	7	100.0	17	6	A65762	A65762 Sequence 43

#### ALIGNMENTS

RESULT 1  
AR062865/c 10 bp DNA PAT 29-SEP-1999  
LOCUS AR062865 Sequence 13 from patent US 5843767.  
DEFINITION AR062865  
ACCESSION AR062865  
VERSION AR062865.1 GI:5990556  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 10)  
AUTHORS Beattie,K.L.  
TITLE Microfabricated, flowthrough porous apparatus for discrete  
detection of binding reactions  
JOURNAL Patent: US 5843767-A 13 01-DEC-1998;  
FEATURES Location/Qualifiers  
source 1..10  
/organism="unknown"

BASE COUNT 4 a 4 c 0 g 2 t  
ORIGIN

Query Match 100.0%; Score 7; DB 6; Length 10;  
Best Local Similarity 100.0%; Pred. No. 5.1e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7  
|||||  
Db 7 TGTGGA 1

## RESULT 2

AX073580/c 12 bp DNA PAT 06-FEB-2001  
LOCUS Sequence 2 from Patent WO0104320.  
DEFINITION AX073580  
ACCESSION AX073580  
VERSION AX073580.1 GI:12710003  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM synthetic construct.  
REFERENCE 1 (bases 1 to 12)  
AUTHORS Schmidt,A.C., Skiadopoulos,M.H., Collins,P.L., Murphy,B.R.,  
TITLE Bally,J.E. and Durbin,A.P.  
JOURNAL Attenuated human-bovine chimeric parainfluenza virus (piv) vaccines  
PATENT: WO 0104320-A 2 18-JAN-2001;  
THE GOVERNMENT OF THE UNITED STATES OF AMERICA (US)

FEATURES  
source 1..12  
Location/Qualifiers  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="Sequence flanking site for introduction of Sgr A1  
site for BP1V3 Ka"

BASE COUNT 4 a 4 c 1 g 3 t  
ORIGIN

Query Match 100.0%; Score 7; DB 6; Length 12;  
Best Local Similarity 100.0%; Pred. No. 5e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7  
|||||  
Db 7 TGTGGA 1

RESULT 3  
A00432/c 13 bp DNA PAT 11-FEB-1993  
LOCUS A00432  
DEFINITION Nucleotide sequence 7 from patent number WO9010459.  
ACCESSION A00432  
VERSION A00432.1 GI:14504  
KEYWORDS  
SOURCE unidentified.  
ORGANISM unidentified.  
REFERENCE 1 (bases 1 to 13)  
AUTHORS  
TITLE PHARMACEUTICAL COMPOSITION USEFUL IN THE PREVENTION OR TREATMENT OF  
JOURNAL PAPILLOMAVIRUS-INDUCED TUMOURS  
PATENT: WO 9010459-A 7 20-SEP-1990;  
TRANSgene S.A

FEATURES  
source 1..13  
Location/Qualifiers  
/organism="unidentified"  
/db\_xref="taxon:32644"

BASE COUNT 4 a 3 c 4 g 2 t  
ORIGIN

Query Match 100.0%; Score 7; DB 6; Length 13;  
Best Local Similarity 100.0%; Pred. No. 5e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7  
|||||  
Db 10 TGTGGA 4

RESULT 4  
A25812/c 14 bp DNA PAT 14-MAR-1995  
LOCUS A25812  
DEFINITION polynucleotide 14C12.  
ACCESSION A25812  
VERSION A25812.1 GI:904780  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM synthetic construct.  
REFERENCE 1 (bases 1 to 14)  
AUTHORS  
JOURNAL Patent: FR 2680520-A 19 26-FEB-1993;  
FEATURES  
source 1..14  
Location/Qualifiers  
/organism="synthetic construct"  
/db\_xref="taxon:32630"

BASE COUNT 4 a 4 c 2 g 4 t  
ORIGIN

Query Match 100.0%; Score 7; DB 6; Length 14;  
Best Local Similarity 100.0%; Pred. No. 5e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7  
|||||  
Db 12 TGTGGA 6

RESULT 5  
A40518 14 bp DNA PAT 05-MAR-1997  
LOCUS A40518  
DEFINITION Sequence 55 from Patent WO9425578.  
ACCESSION A40518  
VERSION A40518.1 GI:2296553  
KEYWORDS  
SOURCE unidentified.  
ORGANISM unidentified.  
REFERENCE 1 (bases 1 to 14)  
AUTHORS  
TITLE ANTISENSE-OLIGONUCLEOTIDES FOR THE TREATMENT OF IMMUNOSUPPRESSIVE  
JOURNAL EFFECTS OF TRANSFORMING GROWTH FACTOR--9(b) (TGF--9(b))  
PATENT: WO 9425578-A 55 10-NOV-1994;  
BIOONOSTIK GES (DE)

FEATURES  
source 1..14  
Location/Qualifiers  
/organism="unidentified"  
/db\_xref="taxon:32644"

BASE COUNT 3 a 3 c 4 g 4 t  
ORIGIN

Query Match 100.0%; Score 7; DB 6; Length 14;  
Best Local Similarity 100.0%; Pred. No. 5e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7  
|||||  
Db 3 TGTGGA 9

RESULT 6  
A88271/c 14 bp DNA PAT 22-JAN-2000  
LOCUS A88271  
DEFINITION Sequence 419 from Patent WO9833904.



ACCESSION A88271 GI:6736841  
VERSION A88271.1  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE 1 (bases 1 to 14)  
AUTHORS Brysch,W. and Schlingensiepen,K.  
TITLE AN ANTISENSE OLIGONUCLEOTIDE PREPARATION METHOD  
JOURNAL Patent: WO 9833904-A 419 06-AUG-1998;  
BIOGOSTIK GES (DE); BRYSCH WOLFGANG (DE)  
FEATURES  
source 1..14  
/organism="unidentified"  
/db\_xref="taxon:32644"  
BASE COUNT 5 a 4 c 0 g 5 t  
ORIGIN

Query Match 100.0%; Score 7; DB 6; Length 14;  
Best Local Similarity 100.0%; Pred. No. 5e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7  
|||||  
Db 11 TGTGGA 5

RESULT 7  
A88273/c 14 bp DNA PAT 22-JAN-2000  
LOCUS A88273 Sequence 421 from Patent WO9833904.  
DEFINITION A88273  
ACCESSION A88273.1 GI:6736843  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE 1 (bases 1 to 14)  
AUTHORS Brysch,W. and Schlingensiepen,K.  
TITLE AN ANTISENSE OLIGONUCLEOTIDE PREPARATION METHOD  
JOURNAL Patent: WO 9833904-A 421 06-AUG-1998;  
BIOGOSTIK GES (DE); BRYSCH WOLFGANG (DE)  
FEATURES  
source 1..14  
/organism="unidentified"  
/db\_xref="taxon:32644"  
BASE COUNT 6 a 4 c 0 g 4 t  
ORIGIN

Query Match 100.0%; Score 7; DB 6; Length 14;  
Best Local Similarity 100.0%; Pred. No. 5e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7  
|||||  
Db 12 TGTGGA 6

RESULT 8  
A89045 14 bp DNA PAT 22-JAN-2000  
LOCUS A89045 Sequence 1193 from Patent WO9833904.  
DEFINITION A89045  
ACCESSION A89045.1 GI:6737615  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE 1 (bases 1 to 14)  
AUTHORS Brysch,W. and Schlingensiepen,K.  
TITLE AN ANTISENSE OLIGONUCLEOTIDE PREPARATION METHOD

JOURNAL Patent: WO 9833904-A 1193 06-AUG-1998;  
BIOGOSTIK GES (DE); BRYSCH WOLFGANG (DE)  
FEATURES  
source 1..14  
/organism="unidentified"  
/db\_xref="taxon:32644"  
BASE COUNT 3 a 3 c 4 g 4 t  
ORIGIN

Query Match 100.0%; Score 7; DB 6; Length 14;  
Best Local Similarity 100.0%; Pred. No. 5e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7  
|||||  
Db 3 TGTGGA 9

RESULT 9  
A90238/c 14 bp DNA PAT 22-JAN-2000  
LOCUS A90238 Sequence 419 from Patent EP0856579.  
DEFINITION A90238  
ACCESSION A90238.1 GI:6738752  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE 1 (bases 1 to 14)  
AUTHORS Brysch,W.D. and Schlingensiepen,K.D.  
TITLE An antisense oligonucleotide preparation method  
JOURNAL Patent: EP 0856579-A 419 05-AUG-1998;  
BIOGOSTIK GES (DE)  
FEATURES  
source 1..14  
/organism="unidentified"  
/db\_xref="taxon:32644"  
BASE COUNT 5 a 4 c 0 g 5 t  
ORIGIN

Query Match 100.0%; Score 7; DB 6; Length 14;  
Best Local Similarity 100.0%; Pred. No. 5e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7  
|||||  
Db 11 TGTGGA 5

RESULT 10  
A90240/c 14 bp DNA PAT 22-JAN-2000  
LOCUS A90240 Sequence 421 from Patent EP0856579.  
DEFINITION A90240  
ACCESSION A90240.1 GI:6738754  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE 1 (bases 1 to 14)  
AUTHORS Brysch,W.D. and Schlingensiepen,K.D.  
TITLE An antisense oligonucleotide preparation method  
JOURNAL Patent: EP 0856579-A 421 05-AUG-1998;  
BIOGOSTIK GES (DE)  
FEATURES  
source 1..14  
/organism="unidentified"  
/db\_xref="taxon:32644"  
BASE COUNT 6 a 4 c 0 g 4 t  
ORIGIN

Query Match 100.0%; Score 7; DB 6; Length 14;  
Best Local Similarity 100.0%; Pred. No. 5e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7  
|||||  
Db 12 TGTGGA 6

## RESULT 11

AR119028 AR119028 14 bp DNA PAT 16-MAY-2001  
LOCUS Sequence 154 from patent US 6150092.  
DEFINITION AR119028  
ACCESSION AR119028  
VERSION AR119028.1 GI:14100938  
KEYWORDS  
SOURCE  
ORGANISM Unknown.  
REFERENCE Unclassified.  
1 (bases 1 to 14)  
AUTHORS Uchida, K., Uchida, T., Tanaka, Y., Matsuda, Y. and Kondo, S.  
TITLE Antisense nucleic acid compound targeted to VEGF  
JOURNAL Patent: US 6150092-A 154 21-NOV-2000;  
FEATURES Location/Qualifiers  
source 1..14  
/organism="unknown"

BASE COUNT 2 a 1 c 6 g 5 t  
ORIGIN

Query Match 100.0%; Score 7; DB 6; Length 14;  
Best Local Similarity 100.0%; Pred. No. 5e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7  
|||||  
Db 7 TGTGGA 13

RESULT 12  
AX007733/c AX007733 14 bp DNA PAT 06-SEP-2000  
LOCUS Sequence 275 from Patent WO9967428.  
DEFINITION AX007733  
ACCESSION AX007733  
VERSION AX007733.1 GI:9995430  
KEYWORDS  
SOURCE  
ORGANISM Aids-associated retrovirus.  
Aids-associated retrovirus.  
VIRUSES: Retroviral viruses; Retroviridae.  
REFERENCE 1 (bases 1 to 14)  
AUTHORS Stuyver, L.  
TITLE Method for detection of drug-selected mutations in the hiv protease gene  
JOURNAL Patent: WO 9967428-A 275 29-DEC-1999;  
INNOGENETICS NV (BE); STUYVER LIEVEN (BE)  
FEATURES Location/Qualifiers  
source 1..14  
/organism="Aids-associated retrovirus"  
/db\_xref="taxon:11966"

BASE COUNT 5 a 6 c 0 g 3 t  
ORIGIN

Query Match 100.0%; Score 7; DB 6; Length 14;  
Best Local Similarity 100.0%; Pred. No. 5e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7  
|||||  
Db 12 TGTGGA 6

RESULT 13  
AX007734/c AX007734 14 bp DNA PAT 06-SEP-2000  
LOCUS Sequence 276 from Patent WO9967428.  
DEFINITION AX007734  
ACCESSION AX007734  
VERSION AX007734.1 GI:9995431  
KEYWORDS  
SOURCE  
ORGANISM Aids-associated retrovirus.  
Aids-associated retrovirus.  
VIRUSES: Retroviral viruses; Retroviridae.  
REFERENCE 1 (bases 1 to 14)  
AUTHORS Stuyver, L.  
TITLE Method for detection of drug-selected mutations in the hiv protease gene  
JOURNAL Patent: WO 9967428-A 276 29-DEC-1999;  
INNOGENETICS NV (BE); STUYVER LIEVEN (BE)  
FEATURES Location/Qualifiers  
source 1..14  
/organism="Aids-associated retrovirus"  
/db\_xref="taxon:11966"

BASE COUNT 5 a 6 c 0 g 3 t  
ORIGIN

Query Match 100.0%; Score 7; DB 6; Length 14;  
Best Local Similarity 100.0%; Pred. No. 5e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7  
|||||  
Db 13 TGTGGA 7

RESULT 14  
AX009010 AX009010 14 bp DNA PAT 06-SEP-2000  
LOCUS Sequence 43 from Patent WO9963975.  
DEFINITION AX009010  
ACCESSION AX009010  
VERSION AX009010.1 GI:9996384  
KEYWORDS  
SOURCE  
ORGANISM human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 14)  
AUTHORS Brysch, W., Schlingensiepen, K.H. and Schlingensiepen, R.  
TITLE A method for stimulating the immune system  
JOURNAL Patent: WO 9963975-A 43 16-DEC-1999;  
BIOGENOSIS GES (DE); BRYSCH WOLFGANG (DE); SCHLINGENSIEPEN KARL  
HERMANN (DE); SCHLINGENSIEPEN REIMAR (DE)  
FEATURES Location/Qualifiers  
source 1..14  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

BASE COUNT 3 a 3 c 4 g 4 t  
ORIGIN

Query Match 100.0%; Score 7; DB 6; Length 14;  
Best Local Similarity 100.0%; Pred. No. 5e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7  
|||||  
Db 3 TGTGGA 9

RESULT 15  
I26234/c I26234 14 bp DNA PAT 07-OCT-1996  
LOCUS Sequence 19 from patent US 5556955.  
DEFINITION I26234  
ACCESSION I26234  
VERSION I26234.1 GI:1606104

KEYWORDS

SOURCE

Unknown.

ORGANISM

Unknown.

REFERENCE

1 (bases 1 to 14)

AUTHORS

Verghnaud,G.

TITLE

Process for detection of new polymorphic loci in a DNA sequence, nucleotide sequences forming hybridization probes and their applications

JOURNAL

Patent: US 5556955-A 19 17-SEP-1996;

FEATURES

Location/Qualifiers

BASE COUNT

4 a 4 c 2 g 4 t

ORIGIN

Query Match

100.0%; Score 7; DB 6; Length 14;

Best Local Similarity

100.0%; Pred. No. 5e+05;

Matches

7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY

1 tgttga 7

Db

12 TGTGGA 6

Search completed: December 15, 2001, 03:19:53

Job time: 7683 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 03:29:43 ; Search time 401.91 Seconds  
(Without alignments)  
14.932 Million cell updates/sec

Title: US-09-380-826a-4  
Perfect score: 7  
Sequence: 1 tgttga 7

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

N.Geneseq\_1101.\*  
1: /SIDS2/gcgcdata/geneseq/geneseq/NA1980.DAT.\*  
2: /SIDS2/gcgcdata/geneseq/geneseq/NA1981.DAT.\*  
3: /SIDS2/gcgcdata/geneseq/geneseq/NA1982.DAT.\*  
4: /SIDS2/gcgcdata/geneseq/geneseq/NA1983.DAT.\*  
5: /SIDS2/gcgcdata/geneseq/geneseq/NA1984.DAT.\*  
6: /SIDS2/gcgcdata/geneseq/geneseq/NA1985.DAT.\*  
7: /SIDS2/gcgcdata/geneseq/geneseq/NA1986.DAT.\*  
8: /SIDS2/gcgcdata/geneseq/geneseq/NA1987.DAT.\*  
9: /SIDS2/gcgcdata/geneseq/geneseq/NA1988.DAT.\*  
10: /SIDS2/gcgcdata/geneseq/geneseq/NA1989.DAT.\*  
11: /SIDS2/gcgcdata/geneseq/geneseq/NA1990.DAT.\*  
12: /SIDS2/gcgcdata/geneseq/geneseq/NA1991.DAT.\*  
13: /SIDS2/gcgcdata/geneseq/geneseq/NA1992.DAT.\*  
14: /SIDS2/gcgcdata/geneseq/geneseq/NA1993.DAT.\*  
15: /SIDS2/gcgcdata/geneseq/geneseq/NA1994.DAT.\*  
16: /SIDS2/gcgcdata/geneseq/geneseq/NA1995.DAT.\*  
17: /SIDS2/gcgcdata/geneseq/geneseq/NA1996.DAT.\*  
18: /SIDS2/gcgcdata/geneseq/geneseq/NA1997.DAT.\*  
19: /SIDS2/gcgcdata/geneseq/geneseq/NA1998.DAT.\*  
20: /SIDS2/gcgcdata/geneseq/geneseq/NA1999.DAT.\*  
21: /SIDS2/gcgcdata/geneseq/geneseq/NA2000.DAT.\*  
22: /SIDS2/gcgcdata/geneseq/geneseq/NA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7	100.0	7	19	AAV58898
2	7	100.0	10	20	AAV59938
3	7	100.0	10	21	AAZ78251
4	7	100.0	10	21	AAZ80950
5	7	100.0	10	21	AAZ82650
6	7	100.0	10	21	AAZ84297
7	7	100.0	10	22	AAH18992
8	7	100.0	10	22	AAH35297
9	7	100.0	10	22	AAH42210
10	7	100.0	10	22	AAH43406
11	7	100.0	10	20	AAH7658

C 12	7	100.0	12	22	AAZ84988	BPIV3 nucleotide f
C 13	7	100.0	14	14	AAQ40606	Hypervariable regi
C 14	7	100.0	14	15	AAQ78406	Antisense oligonuc
C 15	7	100.0	14	19	AAV97198	Potato citrate syn
C 16	7	100.0	14	19	AAV48830	ErbB-2 gene antis
C 17	7	100.0	14	19	AAV48832	ErbB-2 gene antis
C 18	7	100.0	14	19	AAV48832	Integrin alpha 6 s
C 19	7	100.0	14	20	AAA21661	Human C-raf target
C 20	7	100.0	14	20	AAV92044	HIV-1 protease gen
C 21	7	100.0	14	21	AAZ97785	HIV-1 protease gen
C 22	7	100.0	14	21	AAZ65483	Immunosuppressant
C 23	7	100.0	14	21	AAZ64788	Substrate for hair
C 24	7	100.0	15	2	AAH10044	Sequence of initia
C 25	7	100.0	15	2	AAH10045	Human CD40 hamme
C 26	7	100.0	15	17	AAZ66645	Mouse B7-2 hamme
C 27	7	100.0	15	17	AAZ66275	Mouse B7-2 hamme
C 28	7	100.0	15	17	AAZ66276	Mouse B7-2 hamme
C 29	7	100.0	15	17	AAZ66274	Mouse B7-2 hamme
C 30	7	100.0	15	17	AAZ66235	Mouse B7-2 hamme
C 31	7	100.0	15	17	AAZ66236	Mouse B7-2 hamme
C 32	7	100.0	15	17	AAZ66237	Mouse B7-2 hamme
C 33	7	100.0	15	17	AAZ66238	Mouse B7-2 hamme
C 34	7	100.0	15	17	AAZ65872	Human B7-2 hamme
C 35	7	100.0	15	17	AAZ65873	Human B7-2 hamme
C 36	7	100.0	15	17	AAZ50248	Rabbit CERP HH rib
C 37	7	100.0	15	17	AAZ50250	Rabbit CERP HH rib
C 38	7	100.0	15	19	AAV49025	rb gene antisense
C 39	7	100.0	15	19	AAV48831	ErbB-2 gene antis
C 40	7	100.0	15	20	AAZ31673	Tag sequence of a
C 41	7	100.0	15	21	AAZ59285	Human NR8 gene pro
C 42	7	100.0	15	21	AAZ59288	Human NR8 gene pro
C 43	7	100.0	15	21	AAZ59298	Human NR8 gene pro
C 44	7	100.0	15	21	AAZ90823	Human NR8 gene pro
C 45	7	100.0	15	21	AAZ90839	Human NR8 gene pro

#### ALIGNMENTS

RESULT 1	
ID AAV58898	standard; DNA: 7 BP.
XX AC AAV58898;	
XX XX	
DT 20-JAN-1999	(first entry)
XX XX	
DE Leptospira nucleotide sequence.	
XX XX	
KW Infection; pathogenic Leptospira; protective immunity; therapy;	
KM diagnosis; ss.	
XX XX	
OS Leptospira sp.	
XX XX	
PN WO9840099-A1.	
XX XX	
PD 17-SEP-1998.	
XX XX	
PF 06-MAR-1998;	98WO-AU00145.
XX XX	
PR 07-MAR-1997;	97AU-0005494.
XX XX	
PA (AGRI-) AGRIC VICTORIA SERVICES PTY LTD.	
PA (PIGR-) PIG RES & DEV CORP.	
XX XX	
PI Chappel RJ;	
XX XX	
DR WPI: 1998-520791/44.	
XX XX	
PT New isolated pathogenic Leptospira bacterium - useful for, e.g	
PT developing products for conferring protective immunity, and for	
PT prophylactic or therapeutic treatment	
XX XX	

PS Claim 15; Page 71; 94pp; English.  
XX  
CC This sequence represents a Leptospira DNA sequence isolated from the  
CC pathogenic Leptospira (US) bacterium of the invention. The bacterium  
CC belongs to serogroup Hurstbridge or serovar hurstbridge or the species  
CC L. falnei. The US bacterium can be used for conferring protective  
CC immunity against pathogenic LS bacteria in humans or animals. The  
CC bacteria can also be used for prophylactic or therapeutic treatment of LS  
CC infections. The DNAs and antibodies may also be used for detection and  
CC diagnosis of past or present LS infection.  
XX  
SQ Sequence 7 BP; 1 A; 0 C; 3 G; 3 T; 0 other;  
  
Query Match 100.0%; Score 7; DB 19; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.2e+08;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 1 tgttga 7  
|||||  
Db 1 tgttga 7  
  
RESULT 2  
AA99938/C  
ID AA99938 standard; DNA; 10 BP.  
XX  
AC AA99938;  
XX  
DT 21-OCT-1999 (first entry)  
XX  
DE Human parkin gene Intron 7 fragment.  
XX  
KM Parkinson's disease related gene; parkin gene; variant; gene therapy;  
XX Intron; ss.  
XX Homo sapiens.  
XX OS  
XX PN WO9940191-A1.  
XX PD 12-AUG-1999.  
XX PF 09-FEB-1999; 99WO-JP00545.  
XX PR 09-FEB-1998; 98JP-0027531.  
XX PA (MIZU/) MIZUNO Y.  
XX PA (SHIM/) SHIMIZU N.  
XX PI Mizuno Y, Shimizu N;  
XX DR WPI: 1999-494295/41.  
XX PT Gene implicated in the pathology of Parkinson's disease, used for  
XX PT treatment of the disease  
XX PS Claim 11; Page 100; 114pp; English.  
XX  
CC This sequence represents a fragment of an intron from the gene of the  
CC invention. The gene has been designated the parkin gene, and variants of  
CC it are implicated in the pathology of Parkinson's disease, and found in  
CC Parkinson's disease patients. The sequences may be used for the  
CC diagnosis, treatment (including gene therapy) and investigation of  
CC Parkinson's disease.  
XX  
SQ Sequence 10 BP; 3 A; 3 C; 1 G; 3 T; 0 other;  
  
Query Match 100.0%; Score 7; DB 20; Length 10;  
Best Local Similarity 100.0%; Pred. No. 3.3e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 1 tgttga 7

Db |||||  
9 TGTGCA 3  
  
RESULT 3  
AA278251  
ID AA278251 standard; DNA; 10 BP.  
XX  
AC AA278251;  
XX  
DT 10-APR-2000 (first entry)  
XX  
DE Human dendritic cell SAGE tag, SEQ ID NO:679.  
XX  
KM SAGE tag: serial analysis of gene expression; antigen-presenting cell;  
KM APC; monocyte-derived dendritic cell; differential gene expression;  
KM immunostimulatory cofactor; costimulatory factor; CTL;  
KM cytotoxic T-lymphocyte; tumour antigen; immunotherapy; anticancer; ss.  
XX  
OS Homo sapiens.  
XX PN WO9965924-A2.  
XX PD 23-DEC-1999.  
XX  
PF 18-JUN-1999; 99WO-US13800.  
XX  
PR 19-JUN-1998; 98US-0089833.  
PR 19-JUN-1998; 98US-0089844.  
PR 19-JUN-1998; 98US-0089853.  
PR 19-JUN-1998; 98US-0089878.  
PR 19-JUN-1998; 98US-0089991.  
PR 19-JUN-1998; 98US-0089992.  
PR 19-JUN-1998; 98US-0089993.  
PR 19-JUN-1998; 98US-0089994.  
PR 19-JUN-1998; 98US-0089997.  
PR 19-JUN-1998; 98US-0089999.  
PR 19-JUN-1998; 98US-0090000.  
PR 19-JUN-1998; 98US-0090035.  
PR 19-JUN-1998; 98US-0090036.  
PR 19-JUN-1998; 98US-0090039.  
PR 19-JUN-1998; 98US-0090040.  
PR 19-JUN-1998; 98US-0090041.  
PR 19-JUN-1998; 98US-0090042.  
PR 19-JUN-1998; 98US-0090043.  
PR 19-JUN-1998; 98US-0090044.  
PR 19-JUN-1998; 98US-0090045.  
PR 19-JUN-1998; 98US-0090047.  
PR 19-JUN-1998; 98US-0090048.  
PR 19-JUN-1998; 98US-0090072.  
PR 19-JUN-1998; 98US-0090076.  
PR 19-JUN-1998; 98US-0090077.  
PR 19-JUN-1998; 98US-0090078.  
PR 19-JUN-1998; 98US-0090079.  
PR 19-JUN-1998; 98US-0090080.  
PR 08-DEC-1998; 98US-0111715.  
XX  
XX (GENZ ) GENZYME CORP.  
XX PA (ROBE/) ROBERTS B L.  
XX PA (SHAN/) SHANKARA S.  
XX PI Roberts BL, Shankara S;  
XX DR WPI: 2000-106077/09.  
XX  
XX Isolated polynucleotides differentially expressed in antigen-presenting  
XX PT cells, useful in gene vaccines against cancer -  
XX  
PS Claim 1; Page 84; 130pp; English.  
XX  
CC Sequences AA277573-279709 represent SAGE (serial analysis of gene  
CC expression) tags used to identify mRNA transcripts encoding  
CC immunostimulatory cofactor proteins which are preferentially or

CC differentially expressed in monocyte-derived dendritic cells compared  
CC with monocytes. Some of the transcripts correspond to known genes or  
CC ESTs (expressed sequence tags) which were previously unknown to be  
CC preferentially or differentially expressed in dendritic cells, while  
CC other transcripts correspond to novel genes. Antigen-presenting cell  
CC (APC)-associated costimulatory factors play an important role in the  
CC activation of the cytotoxic immune response, particularly against tumour  
CC cells. Tumour antigen presentation via the MHC (major histocompatibility  
CC complex) and subsequent recognition by T-cell receptors is alone  
CC insufficient to activate a robust cytotoxic immune response that can  
CC lyse the tumour cells. Immunostimulatory cofactors also being required  
CC for efficient activation of cytotoxic T-lymphocytes (CTLs). Nucleic acid  
CC sequences identified using the SAGE tags have several potential uses.  
CC They may be used in vaccines to induce an immune response, particularly  
CC against a tumour antigen; to modulate the genome type of an APC; to screen  
CC for agents that modulate expression of differentially expressed genes in  
CC an APC; and as hybridisation probes/amplification primers for the  
CC diagnosis, prognosis and monitoring of diseases related to abnormal  
CC expression of these genes. Detection of the dendritic cell  
CC differentially expressed genes, or of their encoded proteins, can be used  
CC to identify cells as belonging to the monocyte lineage. Cells containing  
CC these genes can be used in active immunotherapy (or to stimulate  
CC production of a population of antigen-specific effector cells) and  
CC vectors containing them are used in gene therapy. Co-administration of  
CC tumour antigens and APC-associated costimulatory factors ensures adequate  
CC antigen presentation to endogenous APCs and upregulates the APCs for the  
CC presentation of co-stimulatory signals, migration to T cell-rich sites,  
CC secretion of T cell growth factors and secretion of chemokines for  
CC recruitment of immune effector cells.  
CC  
XX Sequence 10 BP; 1 A; 2 C; 3 G; 4 T; 0 other;  
SQ

Query Match 100.0%; Score 7; DB 21; Length 10;  
Best Local Similarity 100.0%; Pred. No. 3.3e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7  
|||||  
Db 3 tgttga 9

RESULT 4  
AAZ80950  
ID AAZ80950 standard; DNA: 10 BP.  
AC AAZ80950;  
XX  
DT 07-APR-2000 (first entry)  
XX  
DE Metastatic breast tumour cell upregulated transcript tag #184.  
XX  
DE Human; metastatic breast tumour tissue; breast cancer; tag; primer;  
KW non-metastatic breast tumour tissue; gene therapy; anticancer;  
KW antimetastatic; vaccine; diagnosis; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO9965928-A2.  
XX  
PD 23-DEC-1999.  
XX  
PF 18-JUN-1999; 99WO-US13647.  
XX  
PR 19-JUN-1998; 98US-0089853.  
PR 19-JUN-1998; 98US-0089997.  
PR 19-JUN-1998; 98US-0090039.  
PR 19-JUN-1998; 98US-0090040.  
PR 19-JUN-1998; 98US-0090041.  
XX  
PA (GENZ ) GENZYME CORP.  
PA (ROBE/) ROBERTS B L.  
PA (SHAN/) SHANKARA S.

XX  
PI Roberts BL, Shankara S;  
XX  
DR WPI: 2000-106079/09.  
XX  
PT Isolated polynucleotides differentially expressed between metastatic  
PT and non-metastatic breast cancer cells, useful for diagnosis,  
PT prevention and treatment of cancer -  
XX  
PS Claim 1; Page 63; 219pp; English.  
XX

AAZ80767 to AAZ83941 represent tags corresponding to distinct  
CC transcripts that are preferentially transcribed in the metastatic breast  
CC tumour tissue (i.e. are upregulated in metastatic breast tumour cells).  
CC AAZ83942 to AAZ86677 represent tags corresponding to distinct transcripts  
CC that are preferentially transcribed in the primary or non-metastatic  
CC breast tumour tissue (i.e. are downregulated in metastatic breast tumour  
CC cells). These transcripts can be used for diagnosis, prognosis,  
CC monitoring and treatment of breast cancer, particularly where metastatic  
CC diagnosis is by standard immunoassays or hybridisation/amplification  
CC reactions. Compounds that modulate expression of the transcripts are  
CC potentially useful for treatment of (metastatic) breast cancer, while  
CC promoters from the transcripts are used to direct expression, in selected  
CC cell types, of e.g. therapeutic genes (also ribozymes or antisense  
CC sequences), particularly an antigen-encoded sequence for use in gene or  
CC cell-based vaccines. Polypeptides encoded by the transcripts are also  
CC useful in vaccines; for diagnosing breast cancer and for raising  
CC specific antibodies (Ab). Ab are used to detect the polypeptides or as  
CC therapeutic agents. Host cells that produce the polypeptides can be used  
CC to expand and isolate populations of educated, antigen-specific immune  
CC effector cells, e.g. cytotoxic T lymphocytes, and these used for  
CC adoptive immunotherapy.  
XX  
SQ Sequence 10 BP; 2 A; 0 C; 4 G; 4 T; 0 other;

Query Match 100.0%; Score 7; DB 21; Length 10;  
Best Local Similarity 100.0%; Pred. No. 3.3e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7  
|||||  
Db 4 tgttga 10

RESULT 5  
AAZ82660  
ID AAZ82660 standard; DNA: 10 BP.  
AC AAZ82660;  
XX  
DT 07-APR-2000 (first entry)  
XX  
DE Metastatic breast tumour cell upregulated transcript tag #1894.  
XX  
DE Human; metastatic breast tumour tissue; breast cancer; tag; primer;  
KW non-metastatic breast tumour tissue; gene therapy; anticancer;  
KW antimetastatic; vaccine; diagnosis; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO9965928-A2.  
XX  
PD 23-DEC-1999.  
XX  
PF 18-JUN-1999; 99WO-US13647.  
XX  
PR 19-JUN-1998; 98US-0089853.  
PR 19-JUN-1998; 98US-0089997.  
PR 19-JUN-1998; 98US-0090039.  
PR 19-JUN-1998; 98US-0090040.  
PR 19-JUN-1998; 98US-0090041.  
XX

PA (GEN2 ) GENZYME CORP.  
 PA (ROBE/) ROBERTS B L.  
 PA (SHAN/) SHANKARA S.  
 XX  
 PI Roberts BL, Shankara S;  
 XX  
 DR WPI: 2000-106079/09.  
 XX  
 PT Isolated polynucleotides differentially expressed between metastatic  
 PT and non-metastatic breast cancer cells, useful for diagnosis,  
 PT prevention and treatment of cancer -  
 XX  
 PS Claim 1: Page 110; 219pp; English.  
 XX  
 CC AA280767 to AA283941 represent tags corresponding to distinct  
 CC transcripts that are preferentially transcribed in the metastatic breast  
 CC tumour tissue (i.e. are upregulated in metastatic breast tumour cells).  
 CC AA283942 to AA286677 represent tags corresponding to distinct transcripts  
 CC that are preferentially transcribed in the primary or non-metastatic  
 CC breast tumour tissue (i.e. are downregulated in metastatic breast tumour  
 CC cells). These transcripts can be used for diagnosis, prognosis,  
 CC monitoring and treatment of breast cancer, particularly where metastatic.  
 CC diagnosis is by standard immunoassays or hybridisation/amplification  
 CC reactions. Compounds that modulate expression of the transcripts are  
 CC potentially useful for treatment of (metastatic) breast cancer, while  
 CC promoters from the transcripts are used to direct expression, in selected  
 CC cell types, of e.g. therapeutic genes (also ribozymes or antisense  
 CC sequences), particularly an antigen-encoding sequence for use in gene or  
 CC cell-based vaccines. Polypeptides encoded by the transcripts are also  
 CC useful in vaccines; for diagnosing breast cancer and for raising  
 CC specific antibodies (Ab). Ab are used to detect the polypeptides or as  
 CC therapeutic agents. Host cells that produce the polypeptides can be used  
 CC to expand and isolate populations of educated, antigen-specific immune  
 CC effector cells, e.g. cytotoxic T lymphocytes, and these used for  
 CC adoptive immunotherapy.  
 CC  
 SO Sequence 10 BP; 2 A; 0 C; 4 G; 4 T; 0 other;

Query Match 100.0%; Score 7; DB 21; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 3.3e+04;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7  
 |||||  
 Db 2 tgttga 8

RESULT 6  
 AA284297  
 ID AA284297 standard; DNA; 10 BP.  
 XX  
 AC AA284297;  
 XX  
 DT 07-APR-2000 (first entry)  
 XX  
 DE Metastatic breast tumour cell downregulated transcript tag #3531.  
 XX  
 KW Human; metastatic breast tumour tissue; breast cancer; tag; primer;  
 KW non-metastatic breast tumour tissue; gene therapy; anticancer;  
 KW antimetastatic; vaccine; diagnosis; ss.  
 XX  
 OS Homo sapiens.  
 OS  
 PN WO965928-A2.  
 PN  
 PD 23-DEC-1999.  
 PD  
 PF 18-JUN-1999; 99WO-US13647.  
 PF  
 PR 19-JUN-1998; 98US-0089853.  
 PR 19-JUN-1998; 98US-0089997.  
 PR 19-JUN-1998; 98US-0090039.

PR 19-JUN-1998; 98US-0090040.  
 PR 19-JUN-1998; 98US-0090041.  
 XX  
 PA (GEN2 ) GENZYME CORP.  
 PA (ROBE/) ROBERTS B L.  
 PA (SHAN/) SHANKARA S.  
 XX  
 PI Roberts BL, Shankara S;  
 XX  
 DR WPI: 2000-106079/09.  
 XX  
 PT Isolated polynucleotides differentially expressed between metastatic  
 PT and non-metastatic breast cancer cells, useful for diagnosis,  
 PT prevention and treatment of cancer -  
 XX  
 PS Claim 1: Page 153; 219pp; English.  
 XX  
 CC AA280767 to AA283941 represent tags corresponding to distinct  
 CC transcripts that are preferentially transcribed in the metastatic breast  
 CC tumour tissue (i.e. are upregulated in metastatic breast tumour cells).  
 CC AA283942 to AA286677 represent tags corresponding to distinct transcripts  
 CC that are preferentially transcribed in the primary or non-metastatic  
 CC breast tumour tissue (i.e. are downregulated in metastatic breast tumour  
 CC cells). These transcripts can be used for diagnosis, prognosis,  
 CC monitoring and treatment of breast cancer, particularly where metastatic.  
 CC diagnosis is by standard immunoassays or hybridisation/amplification  
 CC reactions. Compounds that modulate expression of the transcripts are  
 CC potentially useful for treatment of (metastatic) breast cancer, while  
 CC promoters from the transcripts are used to direct expression, in selected  
 CC cell types, of e.g. therapeutic genes (also ribozymes or antisense  
 CC sequences), particularly an antigen-encoding sequence for use in gene or  
 CC cell-based vaccines. Polypeptides encoded by the transcripts are also  
 CC useful in vaccines; for diagnosing breast cancer and for raising  
 CC specific antibodies (Ab). Ab are used to detect the polypeptides or as  
 CC therapeutic agents. Host cells that produce the polypeptides can be used  
 CC to expand and isolate populations of educated, antigen-specific immune  
 CC effector cells, e.g. cytotoxic T lymphocytes, and these used for  
 CC adoptive immunotherapy.  
 CC  
 SO Sequence 10 BP; 2 A; 0 C; 3 G; 5 T; 0 other;

Query Match 100.0%; Score 7; DB 21; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 3.3e+04;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7  
 |||||  
 Db 2 tgttga 8

RESULT 7  
 AAH18992  
 ID AAH18992 standard; DNA; 10 BP.  
 XX  
 AC AAH18992;  
 XX  
 DT 21-JUN-2001 (first entry)  
 XX  
 DE UCP3 polymorphism detection allele specific primer #105.  
 XX  
 KW UCP3; uncoupling protein 3; polymorphism; obesity;  
 KW diabetes mellitus; ss.  
 XX  
 OS Homo sapiens.  
 OS  
 PN WO200118232-A2.  
 PN  
 PD 15-MAR-2001.  
 PD  
 PF 08-SEP-2000; 2000WO-US24784.  
 PF  
 PR 08-SEP-1999; 99US-0152789.



XX (GENA-) GENAISSANCE PHARM INC.  
PA (STEP/) STEPHENS J C.  
XX  
PI Chew A, Choi JY, Denton RR, Nandabalan K;  
XX WPI: 2001-218562/22.  
XX  
DR  
XX Nucleic acids encoding uncoupling protein 3 (mitochondrial, proton  
PT carrier) (UCP3) proteins comprising single nucleotide polymorphisms,  
PT useful for the design of drugs for treating obesity -  
PS  
XX Disclosure: Page 23: 94pp: English.  
XX  
CC The present invention relates to the human uncoupling protein 3  
CC (mitochondrial, proton carrier) (UCP3) gene and polymorphisms.  
CC The polymorphisms are associated with obesity, especially  
CC diabetes mellitus associated obesity. They polymorphisms may be  
CC identified and analysed to determine whether an individual is  
CC susceptible to obesity and may be used as the basis for targeted  
CC design of drugs to treat obesity. The present sequence was used in  
CC the identification and amplification of UCP3 polymorphisms.  
XX  
SQ Sequence 10 BP; 2 A; 1 C; 3 G; 4 T; 0 other:  
  
Query Match 100.0%; Score 7; DB 22; Length 10;  
Best Local Similarity 100.0%; Pred. No. 3.3e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 tgttga 7  
Db 1 tgttga 7  
|||||  
1 tgttga 7  
  
RESULT 8  
AAF35297/c  
ID AAF35297 standard; DNA; 10 BP.  
XX  
AC AAF35297;  
XX  
DT 23-MAR-2001 (first entry)  
XX  
DE Yeast NORF gene SAGE tag oligonucleotide SEQ ID NO:2036.  
XX  
KM Yeast; Saccharomyces cerevisiae; characterisation: cell cycle: NORF;  
KW not previously assigned open reading frame; nonannotated ORF; SAGE;  
KW serial analysis of gene expression; antifungal; tag; identification;  
KW linker: PCR primer; ds.  
XX  
OS Saccharomyces cerevisiae.  
XX  
PN WO200077214-A2.  
XX  
PD 21-DEC-2000.  
XX  
PF 14-JUN-2000; 2000WO-US16223.  
XX  
PR 16-JUN-1999; 99US-0335032.  
XX  
PA (UJO ) UNIV JOHNS HOPKINS.  
XX  
PI Velulescu V, Vogelstein B, Kinzler K;  
XX WPI: 2001-061874/07.  
XX  
DR  
XX Yeast gene coding sequences comprising NORF genes with serial analysis  
PT of gene expression (SAGE) tags, useful for studying, monitoring and  
PT affecting phases of the cell cycle -  
PS  
XX Example: Page 72: 419pp: English.  
XX The present invention describes an isolated DNA molecule comprising a

CC coding sequence of a yeast gene selected from a group of 745 NORF (not  
CC previously assigned open reading frame; or nonannotated ORF) genes  
CC comprising a SAGE (serial analysis of gene expression) tag. Also  
CC described are: (1) a method (M1) of using NORF genes to affect the cell  
CC cycle comprising administering a NORF gene whose expression varies by at  
CC least 10% between any two phases of the cell cycle selected from log  
CC phase, S phase and G2/M; (2) a method (M2) for screening candidate  
CC antifungal drugs comprising: (a) contacting a test substance with a  
CC yeast cell; and (b) monitoring expression of a NORF gene whose  
CC expression varies as in M1, where a test substance which modifies the  
CC expression of the yeast gene is a candidate antifungal drug; (3) a method  
CC (M3) for identifying human genes which are involved in cell cycle  
CC progression comprising contacting human DNA with a probe which comprises  
CC at least 10 contiguous nucleotides of a NORF gene whose expression varies  
CC as in M1; and (4) a method (M4) for identifying a candidate drug as a  
CC member of a class of drugs having a characteristic effect on gene  
CC expression in a yeast cell comprising contacting a yeast cell with a  
CC candidate drug and monitoring expression in the yeast cell of at least 1  
CC NORF gene whose expression is affected by the class of drugs. The NORF  
CC genes may be used to study, monitor and affect phases of the cell cycle,  
CC the differentially expressed genes may be used as markers of phases of  
CC the cell cycle. The methods may be used to identify candidate drugs which  
CC affect the cell cycle and for identification of antifungal drugs.  
CC AAF33268 to AAF44064 represent SAGE tags used in the exemplification of  
CC the present invention. AAF33262 to AAF33267 represent linkers and PCR  
CC primers used in the SAGE method, in the exemplification of the present  
CC invention.  
XX  
SQ Sequence 10 BP; 4 A; 4 C; 0 G; 2 T; 0 other:

Query Match 100.0%; Score 7; DB 22; Length 10;  
Best Local Similarity 100.0%; Pred. No. 3.3e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tgttga 7  
Db 10 TGTTCGA 4  
|||||  
10 TGTTCGA 4  
  
RESULT 9  
AAF42210/c  
ID AAF42210 standard; DNA; 10 BP.  
XX  
AC AAF42210;  
XX  
DT 23-MAR-2001 (first entry)  
XX  
DE Yeast NORF gene SAGE tag oligonucleotide SEQ ID NO:8949.  
XX  
KM Yeast; Saccharomyces cerevisiae; characterisation: cell cycle: NORF;  
KW not previously assigned open reading frame; nonannotated ORF; SAGE;  
KW serial analysis of gene expression; antifungal; tag; identification;  
KW linker: PCR primer; ds.  
XX  
OS Saccharomyces cerevisiae.  
XX  
PN WO200077214-A2.  
XX  
PD 21-DEC-2000.  
XX  
PF 14-JUN-2000; 2000WO-US16223.  
XX  
PR 16-JUN-1999; 99US-0335032.  
XX  
PA (UJO ) UNIV JOHNS HOPKINS.  
XX  
PI Velulescu V, Vogelstein B, Kinzler K;  
XX WPI: 2001-061874/07.  
XX  
DR  
XX Yeast gene coding sequences comprising NORF genes with serial analysis  
PT of gene expression (SAGE) tags, useful for studying, monitoring and

PT	affecting phases of the cell cycle
XX	
PS	Example; Page 319; 419pp; English.
VV	

The present invention describes an isolated DNA molecule comprising a coding sequence of a yeast gene selected from a group of 745 NORF (not previously assigned open reading frame; or nonannotated ORF) genes comprising a SAGE (serial analysis of gene expression) tag. Also described are: (1) a method (M1) of using NORF genes to affect the cell cycle comprising administering a NORF gene whose expression varies by at least 10% between any two phases of the cell cycle selected from log phase, S phase and G2/M; (2) a method (M2) for screening candidate antifungal drugs comprising: (a) contacting a test substance with a yeast cell; and (b) monitoring expression of a NORF gene whose expression varies as in M1, where a test substance which modifies the expression of the yeast gene is a candidate antifungal drug; (3) a method (M3) for identifying human genes which are involved in cell cycle progression comprising contacting human DNA with a probe which comprises at least 10 contiguous nucleotides of a NORF gene whose expression varies as in M1; and (4) a method (M4) for identifying a candidate drug as a member of a class of drugs having a characteristic effect on gene expression in a yeast cell comprising contacting a yeast cell with a candidate drug and monitoring expression in the yeast cell of at least 1 NORF gene whose expression is affected by the class of drugs. The NORF genes may be used to study, monitor and affect phases of the cell cycle, the differentially expressed genes may be used as markers of phases of the cell cycle. The methods may be used to identify candidate drugs which affect the cell cycle and for identification of antifungal drugs. AAF33368 to AAF44064 represent SAGE tags used in the exemplification of the present invention. AAF33262 to AAF33267 represent linkers and PCR primers used in the SAGE method, in the exemplification of the present invention.

Sequence 10 BP; 4 A; 3 C; 1 G; 2 T; 0 other;

Query Match	100.0%	Score 7:	DB 22;	Length 10;
Best Local Similarity	100.0%	Pred. No.	3.3e+04;	
Matches	7;	Mismatches	0;	Indels 0; Gaps 0;
		Conservative	0;	

QY	1	tgcttga	7
Db	10	TGTTGGA	4

RESULT	10
AAAF43406	
ID	AAAF43406 standard; DNA; 10. BP.

DT 23-MAR-2001 (first entry)

DE	Yeast NORF	gene SAGE tag	oligonucleotide SEQ	ID NO:11545.
DE				

KM yeast; *Saccharomyces cerevisiae*; characterization; cell cycle; NORF  
 KM nor previously assigned open reading frame; nonannotated ORF; SAGE;  
 KM serial analysis of gene expression; antifungal; tag; identification;  
 KM linker; PCR primer; ds.

OS *Saccharomyces cerevisiae*.

PN WO200077214-A2

PD 21-DEC-2000  
XY

14-JUN-2000; 2000WD-05162233

XX 10-00N-1399; 3905-0333032.

PA (UYJO ) UNIV JOHNS HOPKINS.

PI Velculescu V, Vogelstein B, Kinzler K;

XX WPI; 2001-061874/07.  
DR

PT Yeast gene coding sequences comprising NORF genes with serial analysts  
PT of gene expression (5AGE) tags, useful for studying, monitoring and  
PT affecting phases of the cell cycle -  
XX  
XX Example; Page 362; 419pp; English.

CC The present invention describes an isolated DNA molecule comprising a  
CC coding sequence of a yeast gene selected from a group of 745 NORF (not  
CC previously assigned open reading frame; or nonannotated ORF) genes  
CC comprising a SAGE (serial analysis of gene expression) tag. Also  
CC described are: (1) a method (M1) of using NORF genes to affect the cell  
CC cycle comprising administering a NORF gene whose expression varies by at  
CC least 10% between any two phases of the cell cycle selected from log  
CC phase, S phase and G2/M; (2) a method (M2) for screening candidate  
CC antifungal drugs comprising: (a) contacting a test substance with a  
CC yeast cell; and (b) monitoring expression of a NORF gene whose  
CC expression varies as in M1, where a test substance which modifies the  
CC expression of the yeast gene is a candidate antifungal drug; (3) a method  
CC (M3) for identifying human genes which are involved in cell cycle  
CC progression comprising contacting human DNA with a probe which comprises  
CC at least 10 contiguous nucleotides of a NORF gene whose expression varies  
CC as in M1; and (4) a method (M4) for identifying a candidate drug as a  
CC member of a class of drugs having a characteristic effect on gene  
CC expression in a yeast cell comprising contacting a yeast cell with a  
CC candidate drug and monitoring expression in the yeast cell of at least 1  
CC NORF gene whose expression is affected by the class of drugs. The NORF  
CC genes may be used to study, monitor and affect phases of the cell cycle,  
CC the differentially expressed genes may be used as markers of phases of  
CC the cell cycle. The methods may be used to identify candidate drugs which  
CC affect the cell cycle and for identification of antifungal drugs.  
CC AAF33268 to AAF44064 represent SAGE tags used in the exemplification of  
CC the present invention. AAF33262 to AAF33267 represent linkers and PCR  
CC primers used in the SAGE method, in the exemplification of the present  
CC invention.

Sequence 10 BP; 1 A; 0 C; 6 G; 3 T; 0 other;

Query Match	100.0%	Score 7;	DB 22;	length 10;
Best Local Similarity	100.0%;	Pred. No.	3.3e+04;	
Matches	7;	Conservative	0;	Mismatches 0;
				Indels 0;
				Gaps 0;

QY	1	tglttga	7
Db	1	tglttga	7

RESULT	11
AAx77658/c	
ID	AAx77658 standard; DNA; 11 BP.

AC AAX77658;

DT 09-AUG-1999 (first entry)

DE N11 active EGS 22.

KW External guide sequence; ECS; target mRNA; identification; diagnostic;  
KW inactivation; essential gene; therapy; ss.

synthetic.

PN W09927135-AZ  
XX

XX  
LD  
XX

CC  
CON  
TSS

[illegible]

PR 21-NOV-1997; 97US-0976220.

Mon Dec 17 07:48:26 2001

us-09-380-826a-4.rng

Page 7

```

XX (INNO-) INNOVIR LAB INC.
XX
XX
PI Kindt TJ, Nilsen TW, Robertson HD,
XX
XX WPI: 1999-357853/30.
XX
PT Identifying and inhibiting functional nucleic acid molecules in
PT cells
XX
XX Example 3; Page 28; 58pp; English.
XX
XX This invention describes a novel method allowing essential or functional
CC genes to be rapidly identified and inactivated. The method is able to
CC firstly identify most of the essential genes in an organism (i.e. a
CC bacteria or a eukaryote) needed for survival, and secondly it provides
CC for reducing or inactivating their expression. The method is able to
CC identify functional oligonucleotide molecules able to be used as
CC diagnostic reagents and therapeutics. The method provides a means for
CC identifying essential genes whose sequence is known only as part of a
CC genome with unknown function, as well as a means for identifying
CC functional oligonucleotide molecules. The method involves the use of a
CC nucleic acid molecule comprising (a) a first reporter gene encoding a
CC fusion protein comprising a protein of interest (itself translated from
CC an RNA of interest) and a reporter protein, a second reporter gene
CC encoding a second reporter protein, and (c) a targeting guide encoding a
CC functional oligonucleotide molecule such as an external guide sequence
CC (EGS), a ribozyme or an antisense RNA and targeted to the RNA of interest
CC at a site on the first reporter gene able to encode the RNA of interest.
XX
SQ Sequence 11 BP; 4 A; 4 C; 1 G; 2 T; 0 other;

Query Match      100.0%; Score 7; DB 20; Length 11;
Best Local Similarity 100.0%; Pred. No. 3.3e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7
   |||||
Db 10 TGTGGA 4

RESULT 12
AAC84988/c
ID AAC84988 standard; DNA; 12 BP.
XX
XX AAC84988;
AC
XX 08-MAY-2001 (first entry)
DT
XX
XX BPIV3 nucleotide fragment.
DE
XX
XX Parainfluenza virus; PIV; infectious; human; nucleocapsid protein; BPIV;
KW nucleocapsid phosphoprotein; polymerase protein; human; bovine; HPIV;
KW chimeric; vaccine; immune response; HPIV1; HPIV2; HPIV3; BPIV3; ds.
XX
XX Bovine parainfluenza virus.
OS
XX
XX W0200104320-A1.
PN
XX
XX 18-JAN-2001.
PD
XX
XX 16-JUN-2000; 2000WO-US17066.
PF
XX
XX 09-JUL-1999; 99US-0143134.
PR
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA
XX
XX Schmidt AC, Skladopoulos MH, Collins PL, Murphy BR, Bailly JE;
PI Durbin AP;
XX
XX WPI: 2001-081053/09.
XX

```

```

PT Isolated human-bovine chimeric parainfluenza virus (PIV), useful in an
PT attenuated vaccine to elicits an immune response against one or more
PT virus(es) selected from human PIV1 (HPIV1), HPIV2 and HPIV3 -
XX
XX
XX Example 4; Page 80; 148pp; English.
XX
XX The invention relates to an isolated human-bovine chimeric parainfluenza
CC virus (PIV) that is infectious and attenuated in humans. The virus
CC comprises a major nucleocapsid protein, a nucleocapsid phosphoprotein,
CC a large polymerase protein, and a partial or complete PIV background
CC genome, or antigenome of a human PIV (HPIV) or bovine PIV (BPIV),
CC combined with one or more heterologous gene(s) or genome segment(s) of a
CC different PIV to form a human-bovine chimeric PIV genome or antigenome.
CC The chimeric PIV is useful in attenuated vaccines to elicit immune
CC response against one or more virus(es) selected from HPIV1, HPIV2 and
CC HPIV3 and most preferably against HPIV3. The present sequence represents
CC a BPIV3 parent nucleotide fragment. This sequence is changed in a
CC recombinant BPIV3 for constructing a chimeric PIV.
XX
SQ Sequence 12 BP; 4 A; 4 C; 1 G; 3 T; 0 other;

Query Match      100.0%; Score 7; DB 22; Length 12;
Best Local Similarity 100.0%; Pred. No. 3.3e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7
   |||||
Db 7 TGTGGA 1

RESULT 13
AAQ40606/c
ID AAQ40606 standard; DNA; 14 BP.
XX
XX AAQ40606;
AC
XX 10-AUG-1993 (first entry)
DT
XX
XX Hypervariable region detection probe 14C12.
DE
XX
XX HVR; human; animal; forensic science; paternity testing; diagnosis;
KW animal breeding; hereditary diseases; tumours; allele; loss;
KW chromosomal regions; tumour region identification; ss.
XX
XX Synthetic.
OS
XX
XX FR2680520-A.
PN
XX
XX 26-FEB-1993.
PD
XX
XX 22-AUG-1991; 91FR-0010516.
PE
XX
XX 22-AUG-1991; 91FR-0010516.
PR
XX
XX (ETFR ) EPAT FRANCAIS.
PA
XX
XX Vergnaud G;
PI
XX
XX WPI: 1993-136548/17.
DR
XX
XX Detecting the hypervariable regions of DNA for diagnosing
PT hereditary illnesses and tumours - by hybridising labelled
PT polynucleotides and analysing genomic DNA of individuals which
PT react with restriction fragments
XX
XX
XX Example; Page 13; 46pp; French.
XX
XX The sequence is that of a polynucleotide probe which may be used in
CC the detection of new hypervariable regions (HVR) in a DNA sequence.
CC HVR represent a fingerprint useful in e.g. forensic science.
CC paternity testing, animal breeding, etc. The probe may be used as
CC part of a method for the efficient detection in humans or other

```

CC animals, without the use of mini-satellites or primary enrichment.  
XX  
SQ Sequence 14 BP; 4 A; 4 C; 2 G; 4 T; 0 other;

Query Match 100.0%; Score 7; DB 14; Length 14;  
Best Local Similarity 100.0%; Pred. No. 3.3e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7  
| | | | |  
DB 12 TGTGGA 6

RESULT 14  
AA078406  
ID AA078406 standard; DNA; 14 BP.  
XX  
AC AA078406;  
XX  
DT 27-JUN-1995 (first entry)  
XX

DE Antisense oligonucleotide hybridising to TGF-beta gene.

XX Transforming growth factor beta; TGF-beta; antisense; treatment;  
KM tumour; angiogenesis; breast tumour; neurofibroma; glioma;  
KM glioblastoma; carcinogenesis; carcinoma; oesophagus; oesophageal;  
KM gastric; gut; immunosuppression; oligonucleotide; ss.

XX Synthetic.

OS

XX

PN W09425588-A.

PD 10-NOV-1994.

XX

PF 29-APR-1994; 94NO-EP01362.

XX

PR 30-APR-1993; 93EP-0107089.

PR 13-MAY-1993; 93EP-0107849.

XX

PA (BIOG-) BIOGNOSTIK GES BIOMOLEKULARE DIAGNOSTIK.

XX

PI Bogdahn U, Brysch W, Schlingensiepen G, Schlingensiepen K;

PI Schlingensiepen R;

XX

DR WPI; 1994-358266/44.

XX

PT New transforming growth factor beta anti:sense

PT oligo:nucleotide(s) - for treating immunosuppression, tumours,

XX etc.

PS Claim 6; Page 40; 74pp; English.

XX

CC The antisense oligonucleotides are useful in the treatment of

CC tumours in which expression of TGF-beta is of relevance for

CC pathogenicity and/or inhibition of pathological angiogenesis. They

CC are used especially for the treatment of the immunosuppressive

CC effect of TGF-beta, augmentation of the proliferation of cytotoxic

CC lymphocytes, treatment of endogenous hyperexpression of TGF-beta,

CC treatment of breast tumours, neurofibromas and malignant gliomas,

CC including glioblastomas, treatment and prophylaxis of skin

CC carcinogenesis, and treatment of oesophageal and gastric carcinomas.

CC See AA078352-078488. The sequences given in GENESQ files

CC AA078352-078407 and AA078408 are antisense oligodeoxynucleotides of

CC TGF-beta 1. The sequences given in GENESQ files AA078408-78487 are

CC antisense oligodeoxynucleotides of TGF-beta 2 in the form of

CC phosphorothioate analogues.

CC  
XX  
SQ Sequence 14 BP; 3 A; 3 C; 4 G; 4 T; 0 other;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7  
| | | | |  
DB 3 tgttga 9

RESULT 15  
AAV97198/C  
ID AAV97198 standard; RNA; 14 BP.  
XX  
AC AAV97198;  
XX  
DT 01-MAR-1999 (first entry)  
XX

DE Potato citrate synthase target sequence position 123.

XX Solanidine; glucosyltransferase; potato; citrate synthase; target;  
KM hammerhead ribozyme; hairpin ribozyme; alkaloid biosynthesis;  
KM flower formation; cleavage; solanaceous plant; ss.

OS Solanum tuberosum.

XX

PN W09832843-A2.

PD 30-JUL-1998.

XX

PF 14-JAN-1998; 98MO-US00738.

XX

PR 24-NOV-1997; 97US-0979416.

PR 28-JAN-1997; 97US-0036545.

XX

PA (RIBO-) RIBOZYME PHARM INC.

XX

PI McSwiggen JA, Zwick MG;

PI WPI; 1998-427939/36.

XX

PT New enzymatic nucleic acid(s) - useful for, e.g. reducing alkaloid

PT biosynthesis or regulating flowering

XX

PS Claim 54; Page 59; 79pp; English.

XX

CC The present invention describes enzymatic nucleic acid molecules with

CC RNA-cleaving activity (e.g. ribozymes) which are capable of modulating

CC the expression of plant genes: (i) involved in biosynthesis of

CC alkaloids; or (ii) involved in flower formation. AAV95982 to AAV96334,

CC and AAV96335 to AAV96354 represent potato solanidine glucosyltransferase

CC hammerhead and hairpin ribozymes, respectively. AAV95981,

CC and AAV96355 to AAV96734 represent potato solanidine glucosyltransferase

CC target sequences. AAV96773 to AAV97170, and AAV97171 to AAV97195

CC represent potato citrate synthase hammerhead and hairpin ribozymes,

CC respectively. AAV96735 to AAV96772, and AAV97196 to AAV97220 represent

CC potato citrate synthase target sequences. Ribozymes of the present

CC invention can be used to inhibit the synthesis of toxic alkaloids in

CC solanaceous plants, particularly potato but also tomato, pepper,

CC aubergine and datura or to inhibit flowering in potato, lettuce, spinach,

CC cabbage, brussel sprouts, arugula, kale, collards, chard, beet, turnip,

CC sweet potato and turf grass. Also the ribozymes can be used for RNA

CC manipulation in the same way that restriction endonucleases are for DNA,

CC as well as to examine genetic drift and mutations in plants and to

CC detect specific RNA. The ribozymes can be targeted to specific genes or

CC to consensus sequences within a family of related genes, and being

CC catalytic need to be present at only very low concentrations.

CC  
XX  
SQ Sequence 14 BP; 4 A; 4 C; 5 G; 1 U; 0 other;

Query Match 100.0%; Score 7; DB 19; Length 14;  
Best Local Similarity 100.0%; Pred. No. 3.3e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Mon Dec 17 07:48:26 2001

us-09-380-826a-4.rng

Page 9

OY 1 t9lt9ga 7  
| | | | |  
Db 13 TGTTCGA 7

Search completed: December 15, 2001, 03:29:44  
Job time: 7884 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 03:22:46 : Search time 172.39 Seconds  
(without alignments)  
9.196 Million cell updates/sec

Title: US-09-380-826a-4

Perfect score: 7

Sequence: 1 tgttga 7

Scoring table: IDENTITY\_NUC

Gap 10.0, Capext 1.0

Searched: 351203 seqs, 1133899 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PCBUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed.  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	7	100.0	10	US-08-631-751A-13	Sequence 13, Appl
C 2	7	100.0	11	US-09-136-523-25	Sequence 25, Appl
C 3	7	100.0	14	US-08-303-004-19	Sequence 19, Appl
C 4	7	100.0	14	US-08-765-340-154	Sequence 154, Appl
C 5	7	100.0	15	US-08-182-668A-201	Sequence 201, Appl
C 6	7	100.0	15	US-08-182-668A-246	Sequence 246, Appl
C 7	7	100.0	15	US-08-363-240A-614	Sequence 614, Appl
C 8	7	100.0	15	US-08-363-240A-615	Sequence 615, Appl
C 9	7	100.0	15	US-08-363-240A-616	Sequence 616, Appl
C 10	7	100.0	15	US-08-774-306A-201	Sequence 201, Appl
C 11	7	100.0	15	US-08-774-306A-246	Sequence 246, Appl
C 12	7	100.0	15	US-08-585-684B-1367	Sequence 1367, Appl
C 13	7	100.0	15	US-08-585-684B-1367	Sequence 1367, Appl
C 14	7	100.0	15	US-08-585-684B-1729	Sequence 1729, Appl
C 15	7	100.0	15	US-08-585-684B-1730	Sequence 1730, Appl
C 16	7	100.0	15	US-08-585-684B-1731	Sequence 1731, Appl
C 17	7	100.0	15	US-08-585-684B-1732	Sequence 1732, Appl
C 18	7	100.0	15	US-08-585-684B-1768	Sequence 1768, Appl
C 19	7	100.0	15	US-08-585-684B-1769	Sequence 1769, Appl
C 20	7	100.0	15	US-08-585-684B-1770	Sequence 1770, Appl
C 21	7	100.0	15	US-08-585-684B-2139	Sequence 2139, Appl
C 22	7	100.0	15	US-09-064-156A-201	Sequence 201, Appl
C 23	7	100.0	15	US-09-064-156A-246	Sequence 246, Appl
C 24	7	100.0	15	US-09-038-073-1367	Sequence 1367, Appl
C 25	7	100.0	15	US-09-038-073-1367	Sequence 1367, Appl
C 26	7	100.0	15	US-09-038-073-1729	Sequence 1729, Appl
C 27	7	100.0	15	US-09-038-073-1730	Sequence 1730, Appl

C 28	7	100.0	15	US-09-038-073-1731	Sequence 1731, Appl
C 29	7	100.0	15	US-09-038-073-1732	Sequence 1732, Appl
C 30	7	100.0	15	US-09-038-073-1768	Sequence 1768, Appl
C 31	7	100.0	15	US-09-038-073-1769	Sequence 1769, Appl
C 32	7	100.0	15	US-09-038-073-1770	Sequence 1770, Appl
C 33	7	100.0	15	US-09-038-073-2139	Sequence 2139, Appl
C 34	7	100.0	16	US-08-435-350-12	Sequence 32, Appl
C 35	7	100.0	16	US-08-166-664-15	Sequence 15, Appl
C 36	7	100.0	16	US-08-469-461-15	Sequence 15, Appl
C 37	7	100.0	16	US-07-890-609-15	Sequence 15, Appl
C 38	7	100.0	17	US-08-072-063-4	Sequence 4, Appl
C 39	7	100.0	17	US-08-119-773-23	Sequence 23, Appl
C 40	7	100.0	17	US-08-281-940-25	Sequence 25, Appl
C 41	7	100.0	17	US-08-390-850-433	Sequence 433, Appl
C 42	7	100.0	17	US-08-064-693-4	Sequence 4, Appl
C 43	7	100.0	17	US-08-373-124A-810	Sequence 810, Appl
C 44	7	100.0	17	US-08-441-370-2	Sequence 2, Appl
C 45	7	100.0	17	US-09-282-146-10	Sequence 10, Appl

## ALIGNMENTS

RESULT 1  
US-08-631-751A-13/C  
Sequence 13, Application US/08631751A  
Patent No. 5843767  
GENERAL INFORMATION:  
APPLICANT: Beattie, Kenneth L.  
TITLE OF INVENTION: Microfabricated, Flowthrough Porous Apparatus for Discrete Detection of Binding Reactions  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Vinson & Elkins  
STREET: 1455 Pennsylvania Avenue, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20004-1008  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/631,751A  
FILING DATE: 11-Apr-11-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Sanzo, Michael A.  
REGISTRATION NUMBER: 36,912  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)639-6500  
TELEFAX: (202)639-6604  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 base pairs  
TYPE: nucleic acid  
TOPOLOGY: linear  
US-08-631-751A-13

Query Match 100.0%; Score 7; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 6.7e+03;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 tgttga 7  
|||||  
Db 7 tgttga 1

RESULT 2  
US-09-196-523-25/c  
Sequence 25, Application US/09196523A  
Patent No. 6248525  
GENERAL INFORMATION:  
APPLICANT: Nilsen, Timothy W.  
TITLE OF INVENTION: Method for Identifying and Inactivating Essential or  
FILE REFERENCE: ILI 130  
CURRENT APPLICATION NUMBER: US/09/196,523A  
CURRENT FILING DATE: 1998-11-20  
EARLIER APPLICATION NUMBER: 60/079,851  
EARLIER FILING DATE: 1998-03-30  
NUMBER OF SEQ ID NOS: 52  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 25  
LENGTH: 11  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence:  
US-09-196-523-25

Query Match 100.0%; Score 7; DB 4; Length 11;  
Best Local Similarity 100.0%; Pred. No. 6.7e+03;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tgttga 7  
|||||  
DB 10 TGTGGA 4

RESULT 3  
US-08-303-004-19/c  
Sequence 19, Application US/08303004  
Patent No. 5556955  
GENERAL INFORMATION:  
APPLICANT: Vergnaud, Gilles  
TITLE OF INVENTION: Process for Detection of New Polymor-  
TITLE OF INVENTION: Phic Loci in an ADN Sequence, Nucleotide Sequences Forming  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Oliff & Herridge  
STREET: P.O. Box 19928  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: U.S.A  
ZIP: 22320  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/303,004  
FILING DATE:  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/931,311B  
FILING DATE: 19920818  
ATTORNEY/AGENT INFORMATION:  
NAME: Herridge, William P.  
REGISTRATION NUMBER: 30,024  
REFERENCE/DOCKET NUMBER: WPB 28264  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-6400  
TELEFAX: (703) 836-2787  
TELEX: 90-1799 PTO ALEX  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:

LENGTH: 14 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-303-004-19

Query Match 100.0%; Score 7; DB 1; Length 14;  
Best Local Similarity 100.0%; Pred. No. 6.8e+03;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tgttga 7  
|||||  
DB 12 TGTGGA 6

RESULT 4  
US-08-765-340-154  
Sequence 154, Application US/08765340  
Patent No. 6150092  
GENERAL INFORMATION:  
APPLICANT: UCHIDA, K.  
APPLICANT: UCHIDA, T.  
APPLICANT: TANAKA, Y.  
APPLICANT: MATSUDA, Y.  
APPLICANT: KONDO, S.  
TITLE OF INVENTION: AN ANTISENSE NUCLEIC ACID  
NUMBER OF SEQUENCES: 185  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/765,340  
FILING DATE: 23-DEC-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 145146/94  
FILING DATE: 27-JUN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 31130/94  
FILING DATE: 21-NOV-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: SERUNIAN, LESLIE  
REGISTRATION NUMBER: 35,353  
REFERENCE/DOCKET NUMBER: 1452-4005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 154:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: other nucleic acid  
DESCRIPTION: /desc = "synthetic DNA"  
US-08-765-340-154

Query Match 100.0%; Score 7; DB 3; Length 14;



Best Local Similarity 100.0%; Pred. No. 6.8e+03;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 ttttga 7  
Db 7 ttttga 13

RESULT 5  
US-08-182-968A-201  
Sequence 201, Application US/08182968A  
Patent No. 5610054  
GENERAL INFORMATION:  
APPLICANT: Draper, Kenneth G.  
TITLE OF INVENTION: METHOD AND REAGENT FOR  
TITLE OF INVENTION: INHIBITING HEPATITIS C  
NUMBER OF SEQUENCES: 497  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
STREET: Suite 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071-2066  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/182,968A  
FILING DATE: 13-JANUARY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/882,888  
FILING DATE: 14-MAY-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Waiburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 205/277  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 201:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-182-968A-201

Query Match 100.0%; Score 7; DB 1; Length 15;  
Best Local Similarity 57.1%; Pred. No. 6.8e+03;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 ttttga 7  
Db 5 ucuuuga 11  
RESULT 6  
US-08-182-968A-246/C  
Sequence 246, Application US/08182968A  
Patent No. 5610054  
GENERAL INFORMATION:  
APPLICANT: Draper, Kenneth G.  
TITLE OF INVENTION: METHOD AND REAGENT FOR  
TITLE OF INVENTION: INHIBITING HEPATITIS C  
TITLE OF INVENTION: VIRUS REPLICATION

NUMBER OF SEQUENCES: 497  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
STREET: Suite 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071-2066  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/182,968A  
FILING DATE: 13-JANUARY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/882,888  
FILING DATE: 14-MAY-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Waiburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 205/277  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 246:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-182-968A-246

Query Match 100.0%; Score 7; DB 1; Length 15;  
Best Local Similarity 100.0%; Pred. No. 6.8e+03;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 ttttga 7  
Db 13 ttttga 7

RESULT 7  
US-08-363-240A-614/C  
Sequence 614, Application US/08363240A  
Patent No. 5705388  
GENERAL INFORMATION:  
APPLICANT: Couture, Larry  
APPLICANT: McSwigen, James  
APPLICANT: Bisgaier, Charles  
APPLICANT: Pape, Michael  
TITLE OF INVENTION: METHOD AND REAGENT FOR  
TITLE OF INVENTION: PREVENTION, INHIBITION OF  
TITLE OF INVENTION: PROGRESSION AND REGRESSION  
TITLE OF INVENTION: OF VASCULAR DISEASES  
NUMBER OF SEQUENCES: 1243  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
STREET: Suite 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM Compatible

OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/363,240A  
FILING DATE: December 23, 1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 210/096  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ. ID NO: 614:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-363-240A-614

Query Match 100.0%; Score 7; DB 1; Length 15;  
Best Local Similarity 100.0%; Pred. No. 6.8e+03;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7  
1111111  
Db 15 TGTGGA 9

RESULT 8  
US-08-363-240A-615/C  
Sequence 615, Application US/08363240A  
Patent No. 5705388  
GENERAL INFORMATION:  
APPLICANT: Couture, Larry  
APPLICANT: McSwiggen, James  
APPLICANT: Bisgaler, Charles  
APPLICANT: Pape, Michael  
TITLE OF INVENTION: METHOD AND REAGENT FOR  
PREVENTION, INHIBITION OF  
TITLE OF INVENTION: PROGRESSION AND REGRESSION  
TITLE OF INVENTION: OF VASCULAR DISEASES  
NUMBER OF SEQUENCES: 1243  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
STREET: Suite 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/363,240A  
FILING DATE: December 23, 1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 210/096  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ. ID NO: 615:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-363-240A-615

Query Match 100.0%; Score 7; DB 1; Length 15;  
Best Local Similarity 100.0%; Pred. No. 6.8e+03;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7  
1111111  
Db 15 TGTGGA 9

RESULT 9  
US-08-363-240A-616/C  
Sequence 616, Application US/08363240A  
Patent No. 5705388  
GENERAL INFORMATION:  
APPLICANT: Couture, Larry  
APPLICANT: McSwiggen, James  
APPLICANT: Bisgaler, Charles  
APPLICANT: Pape, Michael  
TITLE OF INVENTION: METHOD AND REAGENT FOR  
PREVENTION, INHIBITION OF  
TITLE OF INVENTION: PROGRESSION AND REGRESSION  
TITLE OF INVENTION: OF VASCULAR DISEASES  
NUMBER OF SEQUENCES: 1243  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
STREET: Suite 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/363,240A  
FILING DATE: December 23, 1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 210/096  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ. ID NO: 616:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-363-240A-616

Query Match 100.0%; Score 7; DB 1; Length 15;

Best Local Similarity 100.0%; Pred. No. 6.8e+03;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 tqtgga 7  
Db 13 ttttga 7

RESULT 10

US-08-774-306A-201  
Sequence 201, Application US/08774306A  
Patent No. 5869253  
GENERAL INFORMATION:  
APPLICANT: Draper, Kenneth G.  
TITLE OF INVENTION: METHOD AND REAGENT FOR  
TITLE OF INVENTION: INHIBITING HEPATITIS C  
NUMBER OF SEQUENCES: 497  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
CITY: Suite 4700  
STATE: Los Angeles  
COUNTRY: U.S.A.  
ZIP: 90071-2066  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/774,306A  
FILING DATE: December 26, 1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/182,968  
FILING DATE: January 13, 1994  
APPLICATION NUMBER: 07/882,888  
FILING DATE: May 14, 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 223/227  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 201:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-774-306A-201

Query Match 100.0%; Score 7; DB 2; Length 15;  
Best Local Similarity 57.1%; Pred. No. 6.8e+03;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 1 tqtgga 7  
Db 5 ucucga 11

RESULT 11

US-08-774-306A-246/C  
Sequence 246, Application US/08774306A  
Patent No. 5869253  
GENERAL INFORMATION:  
APPLICANT: Draper, Kenneth G.  
TITLE OF INVENTION: METHOD AND REAGENT FOR

TITLE OF INVENTION: INHIBITING HEPATITIS C  
TITLE OF INVENTION: VIRUS REPLICATION  
NUMBER OF SEQUENCES: 497  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
CITY: Suite 4700  
STATE: Los Angeles  
COUNTRY: U.S.A.  
ZIP: 90071-2066  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/774,306A  
FILING DATE: December 26, 1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/182,968  
FILING DATE: January 13, 1994  
APPLICATION NUMBER: 07/882,888  
FILING DATE: May 14, 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 223/227  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 246:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-774-306A-246

Query Match 100.0%; Score 7; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 6.8e+03;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 tqtgga 7  
Db 13 ttttga 7

RESULT 12

US-08-585-684B-1366/C  
Sequence 1366, Application US/08585684B  
Patent No. 5877021  
GENERAL INFORMATION:  
APPLICANT: Stinchcomb, Daniel T.  
APPLICANT: McSwigen, James  
TITLE OF INVENTION: METHOD AND REAGENT FOR THE  
TITLE OF INVENTION: INDUCTION OF GRAFT TOLERANCE  
NUMBER OF SEQUENCES: 2751  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
CITY: Suite 4700  
STATE: Los Angeles  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: FASTSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/585,684B  
FILING DATE: January 16, 1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/000,951  
FILING DATE: July 7, 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 218/078  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 1366:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-585-684B-1366

Query Match 100.0%; Score 7; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 6.8e+03;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 tgttga 7  
|||||  
Db 14 TGTTCGA 8

RESULT 13  
US-08-585-684B-1367/c  
Sequence 1367, Application US/08585684B  
Patent No. 5877021  
GENERAL INFORMATION:  
APPLICANT: Stinchcomb, Daniel T.  
APPLICANT: Jarvis, Thale  
APPLICANT: McSwigen, James  
TITLE OF INVENTION: METHOD AND REAGENT FOR THE  
TITLE OF INVENTION: INDUCTION OF GRAFT TOLERANCE  
TITLE OF INVENTION: AND REVERSAL OF IMMUNE RESPONSES  
NUMBER OF SEQUENCES: 2751  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
STREET: Suite 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: FASTSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/585,684B  
FILING DATE: January 16, 1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/000,951  
FILING DATE: July 7, 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 218/078  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 1367:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-585-684B-1367

Query Match 100.0%; Score 7; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 6.8e+03;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 tgttga 7  
|||||  
Db 13 TGTTCGA 7

RESULT 14  
US-08-585-684B-1729/c  
Sequence 1729, Application US/08585684B  
Patent No. 5877021  
GENERAL INFORMATION:  
APPLICANT: Stinchcomb, Daniel T.  
APPLICANT: Jarvis, Thale  
APPLICANT: McSwigen, James  
TITLE OF INVENTION: METHOD AND REAGENT FOR THE  
TITLE OF INVENTION: INDUCTION OF GRAFT TOLERANCE  
TITLE OF INVENTION: AND REVERSAL OF IMMUNE RESPONSES  
NUMBER OF SEQUENCES: 2751  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
STREET: Suite 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: FASTSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/585,684B  
FILING DATE: January 16, 1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/000,951  
FILING DATE: July 7, 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 218/078  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 1729:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-585-684B-1729

Query Match 100.0%; Score 7; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 6.8e+03;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7  
 111111  
 Db 13 TGTGGA 7

RESULT 15  
 US-08-585-684B-1730/c  
 : Sequence 1730, Application US/08585684B  
 : Patent No. 5877021  
 :  
 : GENERAL INFORMATION:  
 : APPLICANT: Stinchcomb, Daniel T.  
 : APPLICANT: Jarvis, Thale  
 : APPLICANT: McSwiggen, James  
 : TITLE OF INVENTION: METHOD AND REAGENT FOR THE  
 : TITLE OF INVENTION: INDUCTION OF CRAFT TOLERANCE  
 : TITLE OF INVENTION: AND REVERSAL OF IMMUNE RESPONSES  
 : NUMBER OF SEQUENCES: 2751  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Lyon & Lyon  
 : STREET: 633 West Fifth Street  
 : STREET: Suite 4700  
 : CITY: Los Angeles  
 : STATE: California  
 : COUNTRY: U.S.A.  
 : ZIP: 90071  
 :  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
 :  
 : COMPUTER: IBM Compatible  
 : OPERATING SYSTEM: IBM P.C. DOS 5.0  
 : SOFTWARE: FastSeq Version 1.5  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/585,684B  
 : FILING DATE: January 16, 1996  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: 60/000,951  
 : FILING DATE: July 7, 1995  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: Warburg, Richard  
 : REGISTRATION NUMBER: 32,327  
 : REFERENCE/DOCKET NUMBER: 218/078  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: (213) 489-1600  
 : TELEFAX: (213) 955-0440  
 : TELEX: 67-3510  
 : INFORMATION FOR SEQ ID NO: 1730:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 15 base pairs  
 : TYPE: nucleic acid  
 : STRANDEDNESS: single  
 : TOPOLOGY: linear  
 : US-08-585-684B-1730

Query Match 100.0%; Score 7; DB 2; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 6.8e+03;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0;  
 Gaps 0;  
 OY 1 tgttga 7  
 111111  
 Db 13 TGTGGA 7

Search completed: December 15, 2001, 03:22:47  
 Job time: 7562 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 02:33:37 ; Search time 4441.54 seconds  
(without alignments)  
16.936 Million cell updates/sec

Title: US-09-380-826a-4  
Perfect score: 7  
Sequence: 1 tglgtga 7

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues  
Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

- 1: em\_estfun:\*
- 2: em\_esthum:\*
- 3: em\_estln:\*
- 4: em\_estom:\*
- 5: em\_estpl:\*
- 6: em\_estba:\*
- 7: em\_estro:\*
- 8: em\_estov:\*
- 9: em\_hlc:\*
- 10: gb\_estl:\*
- 11: gb\_estc2:\*
- 12: gb\_hlc:\*
- 13: gb\_gss:\*
- 14: em\_gss\_fun:\*
- 15: em\_gss\_hum:\*
- 16: em\_gss\_inv:\*
- 17: em\_gss\_pln:\*
- 18: em\_gss\_pro:\*
- 19: em\_gss\_rtd:\*
- 20: em\_gss\_vrt:\*
- 21: em\_gss\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	7	100.0	19	AZ386406	1M0145C22
2	7	100.0	19	AZ994163	AZ994163 2M0279F05
3	7	100.0	20	AZ484701	1M0311C24
4	7	100.0	20	AZ807221	2M0068G20
5	7	100.0	20	BH000596	2M0288N06
6	7	100.0	21	AZ361153	1M0104A17
7	7	100.0	21	AZ450829	AZ450829 1M0249E13
8	7	100.0	21	AZ508369	AZ508369 1M0350A24
9	7	100.0	22	TA175C05P	TA175C05P
10	7	100.0	23	AZ663995	AZ663995 1M0543M22
11	7	100.0	23	AZ819914	AZ819914 2M0091N16
12	7	100.0	24	AZ999558	AZ999558 2M0286N24

13	7	100.0	24	TA143H100	AL467041 T. brucei
14	7	100.0	24	TA206G04P	AL475439 T. brucei
15	7	100.0	25	AI074857	AI074857 oy27b11.s
16	7	100.0	25	AI158948	AI158948 u142e07.x
17	7	100.0	25	AI453394	AI453394 t3j7g02.x
18	7	100.0	25	AZ869367	AZ869367 2M0181P03
19	7	100.0	26	AZ352012	AZ352012 1M0090M13
20	7	100.0	26	AZ637522	AZ637522 1M0496N23
21	7	100.0	27	TA346G04Q	AL496254 T. brucei
22	7	100.0	28	AZ777241	AZ777241 2M0011G02
23	7	100.0	29	AZ332711	AZ332711 1M0061N14
24	7	100.0	29	AZ419519	AZ419519 1M0196N03
25	7	100.0	29	AZ939009	AZ939009 2M0197L13
26	7	100.0	30	AZ580757	AZ580757 1M0369D07
27	7	100.0	30	AZ829102	AZ829102 2M0106K02
28	7	100.0	30	AZ942800	AZ942800 2M0203D12
29	7	100.0	30	TA160D01P	AL473241 T. brucei
30	7	100.0	31	TA251B07P	AL483453 T. brucei
31	7	100.0	31	AZ333172	AZ333172 1M0062B09
32	7	100.0	31	AZ357647	AZ357647 1M0099D19
33	7	100.0	31	TA140G08Q	AL466834 T. brucei
34	7	100.0	32	AV833440	AV833440 AV833440
35	7	100.0	32	AZ320254	AZ320254 1M0040P07
36	7	100.0	32	AZ328463	AZ328463 1M0052D18
37	7	100.0	32	AZ467843	AZ467843 1M0279F06
38	7	100.0	32	AZ587241	AZ587241 1M0394D14
39	7	100.0	32	AZ607412	AZ607412 1M0429A24
40	7	100.0	32	TA187G01Q	AL476327 T. brucei
41	7	100.0	33	H48291	H48291 y977a01.r1
42	7	100.0	34	AZ408169	AZ408169 1M0179C17
43	7	100.0	34	AZ764028	AZ764028 1M0559G13
44	7	100.0	34	AZ775415	AZ775415 2M0007M16
45	7	100.0	34	AZ820498	AZ820498 2M0092C01

# ALIGNMENTS

RESULT 1  
LOCUS AZ386406 19 bp DNA  
DEFINITION 1M0145C22P Mouse 10kb plasmid UGCC1M library Mus musculus genomic clone UGCC1M0145C22 F, DNA sequence.  
ACCESSION AZ386406  
VERSION AZ386406.1 GI:10500106  
KEYWORDS GSS.  
SOURCE house mouse.  
ORGANISM Mus musculus  
REFERENCE 1 (bases 1 to 19)  
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamli,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausern,A. and Wright,D., Weiss,R.  
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts  
JOURNAL Unpublished (2000)  
COMMENT Contact: Robert B. Weiss  
University of Utah Genome Center  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT 84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0145 row: C column: 22  
Seq primer: CGGTGTAAACGACGCCACGT  
Class: plasmid ends  
High quality sequence stop: 19.  
FEATURES  
source 1..19  
Location/Qualifiers

```

/organism="Mus musculus"
/strain="C57BL/6J"
/db.xref="taxon:10090"
/clone="UUGC1M0145C22"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab.host="E. coli strain XL10-gold, T1-resistant, F-"
/notes="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g11473211419b1AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT      5 a      2 c      6 g      6 t
ORIGIN

```

```

Query Match      100.0%; Score 7; DB 13; Length 19;
Best Local Similarity 100.0%; Pred. No. 3.6e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 tgttga 7
    |||||
Db 11 TGTTCGA 17

```

```

RESULT 2
LOCUS   A2994163      19 bp      DNA      GSS      27-APR-2001
DEFINITION 2M0279F05R Mouse 10kb plasmid UUGC2M library Mus musculus genomic
ACCESSION A2994163
VERSION   A2994163.1 GI:13865390
KEYWORDS GSS.
SOURCE    house mouse.
ORGANISM Mus musculus

```

```

REFERENCE 1 (bases 1 to 19)
AUTHORS  Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
          Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
          M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
          and Wright,D.,Weiss,R.

```

```

TITLE    Mouse whole genome scaffolding with paired end reads from 10kb
          plasmid inserts
JOURNAL  Unpublished (2000)
COMMENT  Contact: Robert B. Weiss
          University of Utah Genome Center
          University of Utah
          Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLG, UT
          84112 USA
          Tel: 801 585 5606
          Fax: 801 585 7177
          Email: ddunn@genetics.utah.edu
          Insert Length: 10000 Std Error: 0.00
          Plate: 0279 row: F column: 05
          Seq primer: CACACGAAACAGCTATGACG
          Class: plasmid ends
          High quality sequence stop: 19.
          Location/Qualifiers

```

FEATURES

```

source
1..19
/organism="Mus musculus"
/strain="C57BL/6J"
/db.xref="taxon:10090"
/clone="UUGC2M0279F05"
/clone_lib="Mouse 10kb plasmid UUGC2M library"
/sex="Female"
/lab.host="E. coli strain XL10-gold, T1-resistant, F-"
/notes="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (female) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g11473211419b1AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT      8 a      5 c      3 g      3 t
ORIGIN

```

```

Query Match      100.0%; Score 7; DB 13; Length 19;
Best Local Similarity 100.0%; Pred. No. 3.6e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 tgttga 7
    |||||
Db 11 TGTTCGA 5

```

```

RESULT 3
LOCUS   A2484701      20 bp      DNA      GSS      05-OCT-2000
DEFINITION 1M0311C24F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
ACCESSION A2484701
VERSION   A2484701.1 GI:10649799
KEYWORDS GSS.
SOURCE    house mouse.
ORGANISM Mus musculus

```

```

REFERENCE 1 (bases 1 to 20)
AUTHORS  Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
          Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
          M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
          and Wright,D.,Weiss,R.

```

```

TITLE    Mouse whole genome scaffolding with paired end reads from 10kb
          plasmid inserts
JOURNAL  Unpublished (2000)
COMMENT  Contact: Robert B. Weiss
          University of Utah Genome Center
          University of Utah
          Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLG, UT
          84112, USA
          Tel: 801 585 5606
          Fax: 801 585 7177
          Email: ddunn@genetics.utah.edu
          Insert Length: 10000 Std Error: 0.00
          Plate: 0311 row: C column: 24
          Seq primer: CGTGTAAACGACGCGCAGT
          Class: plasmid ends
          High quality sequence stop: 20.
          Location/Qualifiers

```

FEATURES



## FEATURES

source

Location/Qualifiers

1. .20  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUCG1M0311C24"  
/clone\_1lb="Mouse 10kb plasmid UUCG1M library"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
/note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g114732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT  
ORIGIN  
2 a 4 c 6 g 8 t

Query Match 100.0%; Score 7; DB 13; Length 20;  
Best Local Similarity 100.0%; Pred. No. 3.6e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7  
|||||  
Db 8 TGTTCGA 14

## RESULT 4

A2807221

LOCUS 20 bp DNA GSS 20-FEB-2001

DEFINITION 2M0069G20R Mouse 10kb plasmid UUCG1M library Mus musculus genomic clone UUCG2M0069G20 R, DNA sequence.

ACCESSION A2807221  
VERSION A2807221.1 GI:12971352

KEYWORDS GSS.

SOURCE house mouse.  
ORGANISM Mus musculus

REFERENCE  
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 20)  
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb

JOURNAL  
COMMENT Plasmid inserts  
Unpublished (2000)  
Contact: Robert B. Weiss  
University of Utah Genome Center  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert length: 10000 Std Error: 0.00  
Plate: 0069 row: G column: 20  
Seq primer: CACACAGGAACAGCTATGACC  
Class: plasmid ends

## FEATURES

source

High quality sequence stop: 20.  
Location/Qualifiers

1. .20  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUCG2M0069G20"  
/clone\_1lb="Mouse 10kb plasmid UUCG1M library"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
/note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g114732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT  
ORIGIN  
4 a 2 c 4 g 10 t

Query Match 100.0%; Score 7; DB 13; Length 20;  
Best Local Similarity 100.0%; Pred. No. 3.6e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7  
|||||  
Db 13 TGTTCGA 19

## RESULT 5

BH000596

LOCUS 20 bp DNA GSS 27-APR-2001

DEFINITION 2M0288N06R Mouse 10kb plasmid UUCG2M library Mus musculus genomic clone UUCG2M0288N06 R, DNA sequence.

ACCESSION BH000596  
VERSION BH000596.1 GI:13871822

KEYWORDS GSS.

SOURCE house mouse.  
ORGANISM Mus musculus

REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 20)  
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb

JOURNAL  
COMMENT Plasmid inserts  
Unpublished (2000)  
Contact: Robert B. Weiss  
University of Utah Genome Center  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert length: 10000 Std Error: 0.00  
Plate: 0288 row: N column: 06  
Seq primer: CACACAGGAACAGCTATGACC

Class: plasmid ends  
High quality sequence stop: 20.  
Location/Qualifiers  
1. 20

FEATURES  
source  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUCG2M0288N06"  
/clone\_1lb="Mouse 10kb plasmid UUCG2M library"  
/sex="Female"  
/lab\_host="E. coli strain XL10-Gold, T1-resistant, F-"  
/note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (female) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g1473211419b1AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT  
ORIGIN  
2 a 3 c 8 g 7 t

Query Match  
Best Local Similarity 100.0%; Score 7; DB 13; Length 20;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7  
111111  
Db 1 TGTTCGA 7

RESULT 6  
LOCUS A2361153 21 bp DNA GSS 02-OCT-2000  
DEFINITION 1M0104A17R Mouse 10kb plasmid UUCG1M library Mus musculus genomic  
ACCESSION A2361153  
VERSION A2361153.1 GI:10474853  
KEYWORDS GSS.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 21)  
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausern,A. and Wright,D., Weiss,R.  
Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts  
TITLE Unpublished (2000)  
JOURNAL COMMENT  
Contact: Robert B. Weiss  
University of Utah Genome Center  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0104 row: A column: 17

Seq primer: CACACAGAAACAGCTATGACC  
Class: plasmid ends  
High quality sequence stop: 21.  
Location/Qualifiers  
1. 21

FEATURES  
source  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUCG1M0104A17"  
/clone\_1lb="Mouse 10kb plasmid UUCG1M library"  
/sex="Male"  
/lab\_host="E. coli strain XL10-Gold, T1-resistant, F-"  
/note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g1473211419b1AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT  
ORIGIN  
10 a 5 c 0 g 6 t

Query Match  
Best Local Similarity 100.0%; Score 7; DB 13; Length 21;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7  
111111  
Db 13 TGTTCGA 7

RESULT 7  
LOCUS A2450829 21 bp DNA GSS 04-OCT-2000  
DEFINITION 1M0249E13R Mouse 10kb plasmid UUCG1M library Mus musculus genomic  
ACCESSION A2450829  
VERSION A2450829.1 GI:10606020  
KEYWORDS GSS.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 21)  
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausern,A. and Wright,D., Weiss,R.  
Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts  
TITLE Unpublished (2000)  
JOURNAL COMMENT  
Contact: Robert B. Weiss  
University of Utah Genome Center  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00

Plate: 0249 row: E column: 13  
Seq primer: CACACAGAAACACCTATGACC  
Class: plasmid ends  
High quality sequence stop: 21.  
Location/Qualifiers

FEATURES  
Source

1. .21  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUC1M0249P13"  
/clone\_lib="Mouse 10kb plasmid UUC1M library"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g114732114(gb)AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT  
ORIGIN

8 a 8 c 0 g 5 t

Query Match 100.0%; Score 7; DB 13; Length 21;  
Best Local Similarity 100.0%; Pred. No. 3.6e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 tgttga 7  
|||||  
Db 19 TGTGGA 13

RESULT 8  
LOCUS A2508369 21 bp DNA GSS 05-OCT-2000  
DEFINITION IM0350A24R Mouse 10kb plasmid UUC1M library Mus musculus genomic clone UUC1M0350A24 R, DNA sequence.  
ACCESSION A2508369  
VERSION A2508369.1 GI:10689781  
KEYWORDS GSS.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muride; Murinae; Mus.  
REFERENCE 1 (bases 1 to 21)  
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.  
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts  
JOURNAL Unpublished (2000)  
COMMENT Contact: Robert B. Weiss  
University of Utah Genome Center  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLG, UT 84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunne@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00  
Plate: 0350 row: A column: 24  
Seq primer: CACACAGAAACACCTATGACC  
Class: plasmid ends  
High quality sequence stop: 21.  
Location/Qualifiers

FEATURES  
Source

1. .21  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUC1M0350A24"  
/clone\_lib="Mouse 10kb plasmid UUC1M library"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g114732114(gb)AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT  
ORIGIN

5 a 1 c 7 g 8 t

Query Match 100.0%; Score 7; DB 13; Length 21;  
Best Local Similarity 100.0%; Pred. No. 3.6e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 tgttga 7  
|||||  
Db 3 TGTGGA 9

RESULT 9  
LOCUS TAI75C05P/c 22 bp DNA GSS 13-DEC-2000  
DEFINITION T. brucei sheared genomic DNA clone 175c05, forward sequence, genomic survey sequence.  
ACCESSION AL475173  
VERSION AL475173.1 GI:11840185  
KEYWORDS GSS.  
SOURCE Trypanosoma brucei.  
ORGANISM Trypanosoma brucei.  
Eukaryota; Euzlenozoa; Kinetoplastida; Trypanosomatidae;  
REFERENCE 1 (bases 1 to 22)  
AUTHORS Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R., Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L., Melville,S.E., Rajandream,M.A. and Barrell,B.G.  
TITLE Direct Submission  
JOURNAL Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and nh@sanger.ac.uk  
COMMENT Constructed at the Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from a cloned population of Trypanosoma brucei (TREU927/4 Gnat 10.1) was mechanically sheared to give a tight size distribution (4 kb). The v + i method used for the library construction is described in detail in Smith, H. and Venter, J.C. (Making small

Insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999).  
Email: nelsayed@clif.org  
Details of T. brucei sequencing at the Sanger Centre are available at <http://www.sanger.ac.uk/projects/T-brucei/>.  
Location/Qualifiers

## FEATURES

source

1. 22  
/organism="Trypanosoma brucei"  
/strain="TREU927"  
/db\_xref="taxon:5691"  
/clone="175c05"

BASE COUNT  
ORIGIN

was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g147321141pb|af129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance.

Query Match 100.0%; Score 7; DB 13; Length 22;  
Best Local Similarity 100.0%; Pred. No. 3.6e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tcttga 7  
|||||  
Db 13 TCTTGA 7

RESULT 10  
AZ663995 23 bp DNA GSS 14-DEC-2000  
LOCUS  
DEFINITION  
clone UUGC1M0543M22 R, DNA sequence.

ACCESSION  
AZ663995.1 GI:11801141

VERSION  
KEYWORDS

GSS.

SOURCE

ORGANISM

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 23)

Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,

Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly

,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.

and Wright,D., Weiss,R.

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0543 row: M column: 22

Seq primer: CACACAGGAACACCTATGACC

Class: plasmid ends

High quality sequence stop: 23.

Location/Qualifiers

1. 23

/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="UUGC1M0543M22"

/clone.lib="Mouse 10kb plasmid UUGC1M library"

/sex="Male"

/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"

/note="Vector: PMD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource

(<http://www.jax.org/resources/documents/dnares/>). The DNA

was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA

Query Match 100.0%; Score 7; DB 13; Length 23;  
Best Local Similarity 100.0%; Pred. No. 3.6e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tcttga 7  
|||||  
Db 17 TCTTGA 11

RESULT 11  
AZ819914 23 bp DNA GSS 20-FEB-2001  
LOCUS  
DEFINITION  
clone UUGC2M0091N16 R, DNA sequence.

ACCESSION  
AZ819914

VERSION  
KEYWORDS

GSS.

SOURCE

ORGANISM

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 23)

Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,

Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly

,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.

and Wright,D., Weiss,R.

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0091 row: N column: 16

Seq primer: CACACAGGAACACCTATGACC

Class: plasmid ends

High quality sequence stop: 23.

Location/Qualifiers

1. 23

/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="UUGC2M0091N16"

/clone.lib="Mouse 10kb plasmid UUGC1M library"

/sex="Male"

/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"

/note="Vector: PMD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource

(<http://www.jax.org/resources/documents/dnares/>). The DNA

was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA

0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g114732114|g1b1AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT  
ORIGIN  
5 a 7 c 4 g 7 t

Query Match 100.0%; Score 7; DB 13; Length 23;  
Best Local Similarity 100.0%; Pred. No. 3.6e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 tgttga 7  
|||||  
Db 22 TGTGGA 16

## RESULT 12

A2999558 24 bp DNA GSS 27-APR-2001  
LOCUS 2M0286N24R Mouse 10kb plasmid UUGC2M library Mus musculus genomic  
DEFINITION clone UUGC2M0286N24 R, DNA sequence.

ACCESSION A2999558  
VERSION A2999558.1 GI:13870785  
KEYWORDS GSS.  
SOURCE house mouse.  
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 24) "Barber, M., Beacorn, T., Duval, B., Hamill, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tinney, A., von Niederhausern, A. and Wright, D., Weiss, R.

AUTHORS

Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts  
Unpublished (2000)

JOURNAL COMMENT  
Contact: Robert B. Weiss  
University of Utah Genome Center  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA

Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0286 row: N column: 24  
Seq primer: CACACGACGAACAGCTATGACC  
Class: plasmid ends  
High quality sequence stop: 24.

## FEATURES

source

1. .24

Location/Qualifiers

/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="UUGC2M0286N24"

/clone\_lib="Mouse 10kb plasmid UUGC2M library"

/sex="Female"

/lab\_host="E. coli strain XL10-Gold, T1-resistant, F-"

/note="Vector: pMD42uv; Purified genomic DNA from M. musculus C57BL/6J (female) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g114732114|g1b1AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT  
ORIGIN  
2 a 5 c 8 g 9 t

Query Match 100.0%; Score 7; DB 13; Length 24;  
Best Local Similarity 100.0%; Pred. No. 3.7e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 tgttga 7  
|||||  
Db 17 TGTGGA 23

## RESULT 13

TA143H10Q 24 bp DNA GSS 13-DEC-2000  
LOCUS T. brucei sheared genomic DNA clone 143h10, reverse sequence,  
DEFINITION genomic survey sequence.

ACCESSION A1467041  
VERSION A1467041.1 GI:11836396  
KEYWORDS GSS.  
SOURCE Trypanosoma brucei.  
ORGANISM Trypanosoma brucei

REFERENCE 1 (bases 1 to 24) "Hall, N., Bowman, S., Lennard, N.J., Doggett, J., Atkin, R., Chillingworth, C., Ormond, D., Harris, B., El-Sayed, N., Hou, L., Melville, S.E., Rajadream, M.A. and Barrell, B.G.

AUTHORS

Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and nh@sanger.ac.uk

Constructed at the Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from a cloned population of Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared to give a tight size distribution (4 kb). The v + l method used for the library construction is described in detail in Smith, H. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaubin and B. Barrell, Oxford University Press, 1999).

JOURNAL COMMENT

Email: nelsayed@tigr.org  
Details of T. brucei sequencing at the Sanger Centre are available at [http://www.sanger.ac.uk/Projects/T\\_brucei/](http://www.sanger.ac.uk/Projects/T_brucei/).

## FEATURES

source

1. .24

Location/Qualifiers

/organism="Trypanosoma brucei"

/strain="TREU927"

/db\_xref="taxon:5691"

/clone="143h10"

/clone\_lib="143h10"

/sex="Female"

/lab\_host="E. coli strain XL10-Gold, T1-resistant, F-"

/note="Vector: pMD42uv; Purified genomic DNA from M. musculus C57BL/6J (female) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA

BASE COUNT  
ORIGIN  
2 a 3 c 9 g 10 t

Query Match 100.0%; Score 7; DB 13; Length 24;  
Best Local Similarity 100.0%; Pred. No. 3.7e+05;

```

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 tgttga 7
    1111111
Db 12 TGTGGA 18

RESULT 14
LOCUS TA206G04P/C 24 bp DNA GSS 13-DEC-2000
DEFINITION T. Brucei sheared genomic DNA clone 206g04, forward sequence,
            genomic survey sequence.
ACCESSION AL475439
VERSION AL475439.1 GI:11842420
KEYWORDS GSS.
SOURCE Trypanosoma brucei.
ORGANISM Trypanosoma brucei.
          Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae;
          Trypanosoma.
REFERENCE 1 (bases 1 to 24)
AUTHORS Hall, N., Bowman, S., Lennard, N.J., Doggett, J., Alkin, R.,
          Chillingworth, C., Ormond, D., Harris, B., El-Sayed, N., Hou, L.,
          Melville, S.E., Rajandream, M.A. and Barrell, B.G.
          Direct Submission
          Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
          project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
          Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
          nh@sanger.ac.uk
          Constructed at the Institute for Genomic Research (TIGR),
          Rockville, MD. Genomic DNA isolated from a cloned population of
          Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
          to give a tight size distribution (
          4 kb). The v + i method used for the library construction is
          described in detail in Smith, H. and Venter, J.C. (Making small
          insert libraries for whole genome shotgun sequencing projects. In
          Genome Sequencing: A Practical Approach, eds. M. Vaubin and B.
          Barrell, Oxford University Press, 1999).
          Email: nelsayed@tigr.org
          Details of T. brucei sequencing at the Sanger Centre are available
          at http://www.sanger.ac.uk/projects/T-brucei/.
FEATURES
    source
        1..24
        /organism="Trypanosoma brucei"
        /strain="TREU927"
        /db_xref="taxon:5691"
        /clone="206g04"
BASE COUNT      8 a      8 c      5 g      3 t
ORIGIN
Query Match      100.0%; Score 7; DB 13; Length 24;
Best Local Similarity 100.0%; Pred. No. 3.7e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 tgttga 7
    1111111
Db 11 TGTGGA 5

RESULT 15
LOCUS AI074857 25 bp mRNA EST 24-SEP-1998
DEFINITION oy27b11.s1 Soares-senescent_fibroblasts_NbHSP Homo sapiens cDNA
            clone IMAGE:1667037 3' similar to TR:064507 064507 SERINE 1 ULTRA
            HIGH SULFUR PROTEIN. ; contains PTR5 repetitive element ;, mRNA
            sequence.
ACCESSION AI074857
VERSION AI074857.1 GI:3401501
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

REFERENCE 1 (bases 1 to 25)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
          Tumor Gene Index
          Unpublished (1997)
JOURNAL Contact: Robert Strausberg, Ph.D.
          Email: cgaps-remail.nih.gov
COMMENT This clone is available royalty-free through LNL; contact the
          IMAGE Consortium (info@image.llnl.gov) for further information.
          Trace considered overall poor quality
          Insert Length: 2304 Std Error: 0.00
          Seq primer: -40m13 fwd. ET from Amersham
          High quality sequence stop: 1.
FEATURES
    source
        1..25
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="IMAGE:1667037"
        /clone_lib="Soares-senescent_fibroblasts_NbHSP"
        /tissue_type="senescent_fibroblast"
        /lab_host="DH10B (ampicillin resistant)"
        /note="Vector: pTR73D (Pharmacia) with a modified
          polylinker V-type: phagemid; Site_1: Not I; Site_2: Eco RI
          ; 1st strand cDNA was primed with a Not I - oligo(dT)
          primer [5'
          TCTTACCAATCTGAACTGGAGCGGCCGATTTTTTTTTTTTTTTT 3'],
          double-stranded cDNA was size selected, ligated to Eco RI
          adapters (Pharmacia), digested with Not I and cloned into
          the Not I and Eco RI sites of a modified pTR73 vector
          (Pharmacia). Library went through one round of
          normalization to a Cot = 5. Library constructed by Bento
          Soares and M.Fatima Bonaldo."
BASE COUNT      1 a      4 c      9 g      11 t
ORIGIN
Query Match      100.0%; Score 7; DB 10; Length 25;
Best Local Similarity 100.0%; Pred. No. 3.7e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 tgttga 7
    1111111
Db 2 TGTGGA 8

Search completed: December 15, 2001, 02:33:41
Job time: 4981 sec

```

Mon Dec 17 07:48:29 2001

us-09-380-826a-4.rst

---

THIS PAGE BLANK (USPTO)



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 03:19:53 : Search time 2725.73 Seconds  
(Without alignments)  
42.367 Million cell updates/sec

Title: US-09-380-826a-5  
Perfect score: 1  
Sequence: 1 ttgata 7

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues  
Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl: \*  
1: gb\_da: \*  
2: gb\_htg: \*  
3: gb\_in: \*  
4: gb\_om: \*  
5: gb\_ov: \*  
6: gb\_pat: \*  
7: gb\_ph: \*  
8: gb\_pl: \*  
9: gb\_pr: \*  
10: gb\_ro: \*  
11: gb\_sts: \*  
12: gb\_sy: \*  
13: gb\_un: \*  
14: gb\_vl: \*  
15: em\_ba: \*  
16: em\_fun: \*  
17: em\_hum: \*  
18: em\_in: \*  
19: em\_om: \*  
20: em\_or: \*  
21: em\_ov: \*  
22: em\_pat: \*  
23: em\_ph: \*  
24: em\_pl: \*  
25: em\_ro: \*  
26: em\_sts: \*  
27: em\_sy: \*  
28: em\_un: \*  
29: em\_vl: \*  
30: em\_htgo\_hum: \*  
31: em\_htgo\_inv: \*  
32: em\_htgo\_rod: \*  
33: em\_htg\_hum: \*  
34: em\_htg\_inv: \*  
35: em\_htg\_rod: \*  
36: em\_htg\_other: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
C 1	7	100.0	11 6	120741	120741 Sequence 6
C 2	7	100.0	11 6	120742	120742 Sequence 7
C 3	7	100.0	13 6	AR082373	AR082373 Sequence
C 4	7	100.0	13 6	AR120915	AR120915 Sequence
C 5	7	100.0	13 6	152063	152063 Sequence 5
C 6	7	100.0	13 6	178419	178419 Sequence 21
7	7	100.0	14 6	A88596	A88596 Sequence 74
8	7	100.0	14 6	A90563	A90563 Sequence 74
9	7	100.0	14 6	AR082368	AR082368 Sequence
10	7	100.0	14 6	AR120910	AR120910 Sequence
C 11	7	100.0	14 6	AX007639	AX007639 Sequence
C 12	7	100.0	14 6	178414	178414 Sequence 21
C 13	7	100.0	15 6	A35660	A35660 Synthetic h
C 14	7	100.0	15 6	AR045297	AR045297 Sequence
C 15	7	100.0	15 6	AR071525	AR071525 Sequence
C 16	7	100.0	15 6	AR131688	AR131688 Sequence
17	7	100.0	15 6	AR131689	AR131689 Sequence
18	7	100.0	15 6	AR131690	AR131690 Sequence
19	7	100.0	15 6	AR131691	AR131691 Sequence
C 20	7	100.0	15 6	AR133701	AR133701 Sequence
C 21	7	100.0	15 6	AR133702	AR133702 Sequence
C 22	7	100.0	15 6	AR133703	AR133703 Sequence
C 23	7	100.0	15 6	AR133704	AR133704 Sequence
C 24	7	100.0	15 6	AR133894	AR133894 Sequence
C 25	7	100.0	15 6	AR133895	AR133895 Sequence
C 26	7	100.0	15 6	AR133896	AR133896 Sequence
C 27	7	100.0	15 6	AX007638	AX007638 Sequence
C 28	7	100.0	15 6	AX007639	AX007639 Sequence
C 29	7	100.0	15 6	152349	152349 Sequence 90
30	7	100.0	15 6	161575	161575 Sequence 12
31	7	100.0	15 6	161576	161576 Sequence 13
32	7	100.0	15 6	161577	161577 Sequence 13
33	7	100.0	15 6	177301	177301 Sequence 8
34	7	100.0	15 6	177302	177302 Sequence 9
35	7	100.0	15 6	177303	177303 Sequence 10
C 36	7	100.0	15 6	177948	177948 Sequence 65
C 37	7	100.0	15 6	177949	177949 Sequence 65
C 38	7	100.0	16 6	A64456	A64456 Sequence 7
C 39	7	100.0	16 6	A64457	A64457 Sequence 8
40	7	100.0	16 6	A88876	A88876 Sequence 10
41	7	100.0	16 6	AR141309	AR141309 Sequence
C 42	7	100.0	16 6	AR141310	AR141310 Sequence
C 43	7	100.0	16 6	AX007634	AX007634 Sequence
C 44	7	100.0	16 6	AX007635	AX007635 Sequence
C 45	7	100.0	16 6	AX007637	AX007637 Sequence

## ALIGNMENTS

RESULT	1	PAT	07-OCT-1996
120741/c	120741	11 bp	DNA
LOCUS	Sequence 6 from patent US 5516634.		
DEFINITION	120741		
ACCESSION	120741		
VERSION	120741.1	GT:1601096	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 11)		
AUTHORS	Newman, P.J. and Santoso, S.S.		
TITLE	Molecular basis of the human platelet bra/drba alloantigen system and applications thereof		
JOURNAL	Patent: US 5516634-A 6 14-MAY-1996;		
FEATURES	Location/Qualifiers		
source	1..11		
	/organism="unknown"		
BASE COUNT	5 a	1 c	3 g
ORIGIN			2 t

Query Match 100.0%; Score 7; DB 6; Length 11;  
Best Local Similarity 100.0%; Pred. No. 5.8e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgata 7  
|||||||  
DB 7 TTTGATA 1

RESULT 2  
LOCUS I20742 11 bp DNA PAT 07-OCT-1996  
DEFINITION Sequence 7 from patent US 5516634.  
ACCESSION I20742  
VERSION I20742.1 GI:1601097  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 11)  
AUTHORS Newman,P.J. and Santoso,S.S.  
TITLE Molecular basis of the human platelet bra/brb alloantigen system  
and applications thereof  
JOURNAL Patent: US 5516634-A 7 14-MAY-1996;  
FEATURES Location/Qualifiers  
1..11  
source /organism="unknown"  
BASE COUNT 6 a 1 c 2 g 2 t  
ORIGIN

Query Match 100.0%; Score 7; DB 6; Length 11;  
Best Local Similarity 100.0%; Pred. No. 5.8e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgata 7  
|||||||  
DB 7 TTTGATA 1

RESULT 3  
LOCUS AR082373 13 bp DNA PAT 31-AUG-2000  
DEFINITION Sequence 217 from patent US 5972704.  
ACCESSION AR082373  
VERSION AR082373.1 GI:10009099  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
1 (bases 1 to 13)  
AUTHORS Draper,K.G., Chowrira,B., McSwiggen,J., Stinchcomb,D.T. and  
Thompson,J.D.  
TITLE HIV nef targeted ribozymes  
JOURNAL Patent: US 5972704-A 217 26-OCT-1999;  
FEATURES Location/Qualifiers  
1..13  
source /organism="unknown"  
BASE COUNT 5 a 4 c 1 g 3 t  
ORIGIN

Query Match 100.0%; Score 7; DB 6; Length 13;  
Best Local Similarity 100.0%; Pred. No. 5.7e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgata 7  
|||||||  
DB 10 TTTGATA 4

RESULT 4  
LOCUS AR120915 13 bp DNA PAT 16-MAY-2001  
DEFINITION Sequence 217 from patent US 6159692.  
ACCESSION AR120915  
VERSION AR120915.1 GI:14104491  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
1 (bases 1 to 13)  
AUTHORS Draper,K.G., Chowrira,B., McSwiggen,J., Stinchcomb,D.T. and  
Thompson,J.D.  
TITLE Method and reagent for inhibiting human immunodeficiency virus  
replication  
JOURNAL Patent: US 6159692-A 217 12-DEC-2000;  
FEATURES Location/Qualifiers  
1..13  
source /organism="unknown"

BASE COUNT 5 a 4 c 1 g 3 t  
ORIGIN

Query Match 100.0%; Score 7; DB 6; Length 13;  
Best Local Similarity 100.0%; Pred. No. 5.7e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgata 7  
|||||||  
DB 10 TTTGATA 4

RESULT 5  
LOCUS I52063 13 bp DNA PAT 07-OCT-1997  
DEFINITION Sequence 5 from patent US 5646020.  
ACCESSION I52063  
VERSION I52063.1 GI:2473264  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
1 (bases 1 to 13)  
AUTHORS Swiggen,J.A. and Mamone,J.Anthony.  
TITLE Hammerhead ribozymes for preferred targets  
JOURNAL Patent: US 5646020-A 5 08-JUL-1997;  
FEATURES Location/Qualifiers  
1..13  
source /organism="unknown"  
BASE COUNT 5 a 4 c 1 g 3 t  
ORIGIN

Query Match 100.0%; Score 7; DB 6; Length 13;  
Best Local Similarity 100.0%; Pred. No. 5.7e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgata 7  
|||||||  
DB 10 TTTGATA 4

RESULT 6  
LOCUS I78419 13 bp DNA PAT 03-APR-1998  
DEFINITION Sequence 217 from patent US 5693535.  
ACCESSION I78419  
VERSION I78419.1 GI:3014573  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
1 (bases 1 to 13)

AUTHORS Draper,K.G., Chowrira,B., McSwiggen,J., Stinchcomb,D.T. and Thompson,J.D.  
TITLE HIV targeted ribozymes  
JOURNAL Patent: US 5693535-A 217 02-DEC-1997;  
FEATURES Location/Qualifiers  
source 1..13  
BASE COUNT 5 a 4 c 1 g 3 t  
ORIGIN

Query Match 100.0%; Score 7; DB 6; Length 13;  
Best Local Similarity 100.0%; Pred. No. 5.7e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgata 7  
|||||  
Db 1 TTTGATA 4

RESULT 7  
A88596 14 bp DNA PAT 22-JAN-2000  
LOCUS Sequence 744 from Patent WO9833904.  
ACCESSION A88596  
VERSION A88596.1 GI:6737166  
KEYWORDS  
SOURCE unidentified.  
ORGANISM unidentified  
REFERENCE 1 (bases 1 to 14)  
AUTHORS Brysch,W. and Schlöngensiepen,K.  
TITLE AN ANTISENSE ORIGINUCLEOTIDE PREPARATION METHOD  
JOURNAL Patent: WO 9833904-A 744 06-AUG-1998;  
FEATURES BIOGNOSTIK GES (DE); BRYSCH WOLFGANG (DE)  
source 1..14  
Location/Qualifiers  
/db\_xref="taxon:32644"  
BASE COUNT 5 a 1 c 1 g 7 t  
ORIGIN

Query Match 100.0%; Score 7; DB 6; Length 14;  
Best Local Similarity 100.0%; Pred. No. 5.7e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgata 7  
|||||  
Db 1 TTTGATA 7

RESULT 8  
A90563 14 bp DNA PAT 22-JAN-2000  
LOCUS Sequence 744 from Patent EP0856579.  
ACCESSION A90563  
VERSION A90563.1 GI:6739077  
KEYWORDS  
SOURCE unidentified.  
ORGANISM unidentified  
REFERENCE 1 (bases 1 to 14)  
AUTHORS Brysch,W.D. and Schlöngensiepen,K.D.  
TITLE An antisense oligonucleotide preparation method  
JOURNAL Patent: EP 0856579-A 744 05-AUG-1998;  
FEATURES BIOGNOSTIK GES (DE)  
source 1..14  
Location/Qualifiers  
/db\_xref="taxon:32644"  
BASE COUNT 5 a 1 c 1 g 7 t  
ORIGIN

Query Match 100.0%; Score 7; DB 6; Length 14;  
Best Local Similarity 100.0%; Pred. No. 5.7e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgata 7  
|||||  
Db 1 TTTGATA 7

RESULT 9  
AR082368 14 bp DNA PAT 31-AUG-2000  
LOCUS Sequence 212 from patent US 5972704.  
ACCESSION AR082368  
VERSION AR082368.1 GI:10009094  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 14)  
AUTHORS Draper,K.G., Chowrira,B., McSwiggen,J., Stinchcomb,D.T. and Thompson,J.D.  
TITLE HIV nef targeted ribozymes  
JOURNAL Patent: US 5972704-A 212 26-OCT-1999;  
FEATURES Location/Qualifiers  
source 1..14  
/organism="unknown"  
BASE COUNT 5 a 1 c 4 g 4 t  
ORIGIN

Query Match 100.0%; Score 7; DB 6; Length 14;  
Best Local Similarity 100.0%; Pred. No. 5.7e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgata 7  
|||||  
Db 3 TTTGATA 9

RESULT 10  
ARI20910 14 bp DNA PAT 16-MAY-2001  
LOCUS ARI20910  
DEFINITION Sequence 212 from patent US 6159692.  
ACCESSION ARI20910  
VERSION ARI20910.1 GI:14104486  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 14)  
AUTHORS Draper,K.G., Chowrira,B., McSwiggen,J., Stinchcomb,D.T. and Thompson,J.D.  
TITLE Method and reagent for inhibiting human immunodeficiency virus replication  
JOURNAL Patent: US 6159692-A 212 12-DEC-2000;  
FEATURES Location/Qualifiers  
source 1..14  
/organism="unknown"  
BASE COUNT 5 a 1 c 4 g 4 t  
ORIGIN

Query Match 100.0%; Score 7; DB 6; Length 14;  
Best Local Similarity 100.0%; Pred. No. 5.7e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgata 7  
|||||  
Db 3 TTTGATA 9

```

RESULT 11
LOCUS AX007639 14 bp DNA PAT 06-SEP-2000
DEFINITION Sequence 181 from Patent WO967428.
ACCESSION AX007639
VERSION AX007639.1 GI:9995336
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
BASE COUNT 5 a 1 c 2 g 6 t
ORIGIN

Query Match 100.0%; Score 7; DB 6; Length 14;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ttgata 7
Db 11 TTGATA 5

RESULT 12
LOCUS I78414 14 bp DNA PAT 03-APR-1998
DEFINITION Sequence 212 from patent US 5693535.
ACCESSION I78414
VERSION I78414.1 GI:3014568
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
BASE COUNT 5 a 1 c 4 g 4 t
ORIGIN

Query Match 100.0%; Score 7; DB 6; Length 14;
Best Local Similarity 100.0%; Pred. No. 5.7e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ttgata 7
Db 11 TTGATA 9

RESULT 13
LOCUS A35660 15 bp DNA PAT 02-DEC-1996
DEFINITION Synthetic human IFN-alpha 2 gene oligo.
ACCESSION A35660
VERSION A35660.1 GI:1927042
KEYWORDS
SOURCE
synthetic construct.

```

```

ORGANISM synthetic construct
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
BASE COUNT 5 a 4 c 2 g 4 t
ORIGIN

Query Match 100.0%; Score 7; DB 6; Length 15;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ttgata 7
Db 12 TTGATA 6

RESULT 14
LOCUS AR045297 15 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 90 from patent US 5817796.
ACCESSION AR045297
VERSION AR045297.1 GI:5966762
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
BASE COUNT 6 a 3 c 3 g 3 t
ORIGIN

Query Match 100.0%; Score 7; DB 6; Length 15;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ttgata 7
Db 11 TTGATA 5

RESULT 15
LOCUS AR071525 15 bp DNA PAT 18-FEB-2000
DEFINITION Sequence 25 from patent US 5911982.
ACCESSION AR071525
VERSION AR071525.1 GI:7222413
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
/organism="unknown"

```

BASE COUNT 9 a 2 c 2 g 2 t  
ORIGIN

Query Match 100.0%; Score 7; DB 6; Length 15;  
Best Local Similarity 100.0%; Pred. No. 5.6e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgata 7  
|||  
Db 12 TTTCATA 6

Search completed: December 15, 2001, 03:19:53  
Job time: 7683 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 03:29:44 : Search time 401.91 Seconds  
(Without alignments)  
14.932 Million cell updates/sec

Title: US-09-380-826a-5

Perfect score: 7

Sequence: 1 ttgata 7

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

1: N\_Geneseq\_1101.\*  
2: /SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq/NA1986.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7	100.0	7	19 AAV58899	Leptospira nucleot
2	7	100.0	10	21 AA278577	Human dendritic ce
3	7	100.0	10	21 AA280864	Metastatic breast
4	7	100.0	10	21 AA282567	Metastatic breast
5	7	100.0	10	21 AA282768	Metastatic breast
6	7	100.0	10	21 AA283117	Metastatic breast
7	7	100.0	10	22 AAF34548	Yeast NORF gene SA
8	7	100.0	10	22 AAF35054	Yeast NORF gene SA
9	7	100.0	10	22 AAF38092	Yeast NORF gene SA
10	7	100.0	10	22 AAF39829	Yeast NORF gene SA
11	7	100.0	10	22 AAF40369	Yeast NORF gene SA

12	7	100.0	10	22 AAF40389	Yeast NORF gene SA
13	7	100.0	10	22 AAF40414	Yeast NORF gene SA
14	7	100.0	10	22 AAF40778	Yeast NORF gene SA
15	7	100.0	10	22 AAF41077	Yeast NORF gene SA
16	7	100.0	10	22 AAF41323	Yeast NORF gene SA
17	7	100.0	10	22 AAF42563	Yeast NORF gene SA
18	7	100.0	10	22 AAF43524	Yeast NORF gene SA
19	7	100.0	10	22 AAF43525	Yeast NORF gene SA
20	7	100.0	10	22 AAF43526	Yeast NORF gene SA
21	7	100.0	10	22 AAF43959	Yeast NORF gene SA
22	7	100.0	10	22 AAF44031	Yeast NORF gene SA
23	7	100.0	13	21 AA297719	Hiv-1 protease gen
24	7	100.0	14	19 AAV49155	rb gene antisense
25	7	100.0	14	21 AA297691	HIV-1 protease gen
26	7	100.0	15	14 AA052243	Neuroblastoma spec
27	7	100.0	15	16 AAT56960	RSV IC hammerhead
28	7	100.0	15	16 AAT56962	RSV IC hammerhead
29	7	100.0	15	16 AAT56964	RSV N hammerhead
30	7	100.0	15	16 AAT57424	RSV N hammerhead
31	7	100.0	15	16 AAT57425	Mouse reIA hammer
32	7	100.0	15	16 AAT54843	Mouse reIA hammer
33	7	100.0	15	16 AAT54845	Mouse reIA hammer
34	7	100.0	15	16 AAT54841	Mouse reIA hammer
35	7	100.0	15	17 AAX66632	Human CD40 hamme
36	7	100.0	15	17 AAX66633	Human CD40 hamme
37	7	100.0	15	17 AAX66634	Human CD40 hamme
38	7	100.0	15	17 AAX66635	Human CD40 hamme
39	7	100.0	15	17 AAX66825	Mouse CD40 hamme
40	7	100.0	15	17 AAX66826	Mouse CD40 hamme
41	7	100.0	15	17 AAX66827	Mouse B7-1 hamme
42	7	100.0	15	17 AAX64619	Human B7-1 hamme
43	7	100.0	15	17 AAX64620	Human B7-1 hamme
44	7	100.0	15	17 AAX64621	Human B7-1 hamme
45	7	100.0	15	17 AAX64622	Human B7-1 hamme

## ALIGNMENTS

RESULT 1

AAV58899 standard; DNA; 7 BP.

AAV58899;

20-JAN-1999 (first entry)

Leptospira nucleotide sequence.

Infection: pathogenic Leptospira; protective immunity; therapy;

diagnosis; ss.

Leptospira sp.

WO9840099-A1.

17-SEP-1998.

06-MAR-1998; 98WO-AU00145.

07-MAR-1997; 97AU-0005494.

(AGRI-) AGRIC VICTORIA SERVICES PTY LTD.

(PIGR-) PIG RES & DEV CORP.

Chappel RJ;

WPI; 1998-520791/44.

New isolated pathogenic Leptospira bacterium - useful for, e.g

developing products for conferring protective immunity, and for

prophylactic or therapeutic treatment

PS Claim 15; Page 72; 94pp; English.  
 XX  
 CC This sequence represents a Leptospira DNA sequence isolated from the  
 CC pathogenic leptospira (LS) bacterium of the invention. The bacterium  
 CC belongs to serogroup Hurstbridge or serovar hurstbridge of the species  
 CC L. falnei. The LS bacteria can be used for conferring protective  
 CC immunity against pathogenic LS bacteria in humans or animals. The  
 CC bacteria can also be used for prophylactic or therapeutic treatment of LS  
 CC infections. The DNAs and antibodies may also be used for detection and  
 CC diagnosis of past or present LS infection.  
 XX  
 SO Sequence 7 BP; 2 A; 0 C; 1 G; 4 T; 0 other;  
 Query Match 100.0%; Score 7; DB 19; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+08;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ttgata 7  
 1 tttgata 7  
 Db 1 ttgata 7  
 RESULT 2  
 AA278577/c  
 ID AA278577 standard; DNA; 10 BP.  
 XX  
 AC AA278577;  
 XX  
 DT 10-APR-2000 (first entry)  
 XX  
 DE Human dendritic cell SAGE tag, SEQ ID NO:1005.  
 XX  
 XX SAGE tag; serial analysis of gene expression; antigen-presenting cell;  
 KW APC; monocyte-derived dendritic cell; differential gene expression;  
 KM immunostimulatory cofactor; costimulatory factor; CTL;  
 KM cytotoxic T-lymphocyte; tumour antigen; immunotherapy; anticancer; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN MO965924-A2.  
 XX  
 PD 23-DEC-1999.  
 XX  
 PF 18-JUN-1999; 99MO-US13800.  
 XX  
 XX 19-JUN-1998; 98US-0089833.  
 PR 19-JUN-1998; 98US-0089844.  
 PR 19-JUN-1998; 98US-0089853.  
 PR 19-JUN-1998; 98US-0089878.  
 PR 19-JUN-1998; 98US-0089911.  
 PR 19-JUN-1998; 98US-0089922.  
 PR 19-JUN-1998; 98US-0089932.  
 PR 19-JUN-1998; 98US-0089954.  
 PR 19-JUN-1998; 98US-0089977.  
 PR 19-JUN-1998; 98US-0089997.  
 PR 19-JUN-1998; 98US-0090000.  
 PR 19-JUN-1998; 98US-0090035.  
 PR 19-JUN-1998; 98US-0090036.  
 PR 19-JUN-1998; 98US-0090039.  
 PR 19-JUN-1998; 98US-0090040.  
 PR 19-JUN-1998; 98US-0090041.  
 PR 19-JUN-1998; 98US-0090042.  
 PR 19-JUN-1998; 98US-0090043.  
 PR 19-JUN-1998; 98US-0090044.  
 PR 19-JUN-1998; 98US-0090045.  
 PR 19-JUN-1998; 98US-0090047.  
 PR 19-JUN-1998; 98US-0090048.  
 PR 19-JUN-1998; 98US-0090072.  
 PR 19-JUN-1998; 98US-0090076.  
 PR 19-JUN-1998; 98US-0090077.  
 PR 19-JUN-1998; 98US-0090078.  
 PR 19-JUN-1998; 98US-0090079.

PR 19-JUN-1998; 98US-0090080.  
 PR 08-DEC-1998; 98US-0111715.  
 XX  
 XX (GENZ ) GENZYME CORP.  
 PA (ROBE/) ROBERTS B L.  
 PA (SHAN/) SHANKARA S.  
 XX  
 XX Roberts BL, Shankara S;  
 PI  
 DR WPI, 2000-106077/09.  
 XX  
 XX Isolated polynucleotides differentially expressed in antigen-presenting  
 PT cells, useful in gene vaccines against cancer -  
 PR  
 PS Claim 1; Page 94; 130pp; English.  
 XX  
 XX Sequences AA277573-279709 represent SAGE (serial analysis of gene  
 CC expression) tags used to identify mRNA transcripts encoding  
 CC immunostimulatory cofactor proteins which are preferentially or  
 CC differentially expressed in monocyte-derived dendritic cells compared  
 CC with monocytes. Some of the transcripts correspond to known genes or  
 CC ESTs (expressed sequence tags) which were previously unknown to be  
 CC preferentially or differentially expressed in dendritic cells, while  
 CC other transcripts correspond to novel genes. Antigen-presenting cell  
 CC (APC)-associated costimulatory factors play an important role in the  
 CC activation of the cytotoxic immune response, particularly against tumour  
 CC cells. Tumour antigen presentation via the MHC (major histocompatibility  
 CC complex) and subsequent recognition by T-cell receptors is alone  
 CC insufficient to activate a robust cytotoxic immune response that can  
 CC lyse the tumour cells, immunostimulatory cofactors also being required  
 CC for efficient activation of cytotoxic T-lymphocytes (CTLs). Nucleic acid  
 CC sequences identified using the SAGE tags have several potential uses.  
 CC They may be used in vaccines to induce an immune response, particularly  
 CC against a tumour antigen; to modulate the genotype of an APC; to screen  
 CC for agents that modulate expression of differentially expressed genes in  
 CC an APC; and as hybridisation probes/amplification primers for the  
 CC diagnosis, prognosis and monitoring of diseases related to abnormal  
 CC expression of these genes. Detection of the dendritic cell  
 CC differentially expressed genes, or of their encoded proteins, can be used  
 CC to identify cells as belonging to the monocyte lineage. Cells containing  
 CC these genes can be used in active immunotherapy (or to stimulate  
 CC production of a population of antigen-specific effector cells) and  
 CC vectors containing them are used in gene therapy. Co-administration of  
 CC tumour antigens and APC-associated costimulatory factors ensures adequate  
 CC antigen presentation to endogenous APCs and upregulates the APCs for the  
 CC presentation of co-stimulatory signals, migration to T cell-rich sites,  
 CC secretion of T cell growth factors and secretion of chemokines for  
 CC recruitment of immune effector cells.  
 CC  
 XX  
 SO Sequence 10 BP; 6 A; 1 C; 1 G; 2 T; 0 other;  
 Query Match 100.0%; Score 7; DB 21; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 4.3e+04;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ttgata 7  
 9 tttgata 3  
 Db 9 tttgata 3  
 RESULT 3  
 AA280864  
 ID AA280864 standard; DNA; 10 BP.  
 XX  
 AC AA280864;  
 XX  
 DT 07-APR-2000 (first entry)  
 XX  
 DE Metastatic breast tumour cell upregulated transcript tag #98.  
 DE Metastatic breast tumour tissue; breast cancer; tag; primer;  
 KW Human; metastatic breast tumour tissue; gene therapy; anticancer;  
 KW non-metastatic breast tumour tissue; gene therapy; anticancer;



KW		antimetastatic; vaccine; diagnosis; ss.
XX		Homo sapiens.
OS		
XX		MO965928-A2.
PN		
XX		23-DEC-1999.
PD		
XX		18-JUN-1999;
Pf		99WO-US13647.
XX		
PR		19-JUN-1998;
PR		98US-0089853.
PR		19-JUN-1998;
PR		98US-0089997.
PR		19-JUN-1998;
PR		98US-0090039.
PR		19-JUN-1998;
PR		98US-0090040.
PA		19-JUN-1998;
PA		98US-0090041.
PS	(GENZ ) GENZYME CORP.	
XX	(ROBEF) ROBERTS B L.	
XX	(SHAN/) SHANKARA S.	
PI		
XX	Roberts BL, Shankara S:	
DR		
XX		WPI: 2000-106079/09.
PT		
PT		Isolated polynucleotides differentially expressed between metastatic and non-metastatic breast cancer cells, useful for diagnosis, prevention and treatment of cancer -
PS		
XX	Claim 1; Page 60; 21pp: English.	
CC		
AA80767	to AA83941 represent tags corresponding to distinct	
CC	transcripts that are preferentially transcribed in the metastatic breast	
CC	tumour tissue (i.e. are upregulated in metastatic breast tumour cells).	
CC	AA83942 to AA86677 represent tags corresponding to distinct transcripts	
CC	that are preferentially transcribed in the primary or non-metastatic	
CC	breast tumour tissue (i.e. are downregulated in metastatic breast tumour	
CC	cells). These transcripts can be used for diagnosis, prognosis,	
CC	monitoring and treatment of breast cancer, particularly where metastatic.	
CC	Diagnosis is by standard immunoassays or hybridisation/amplification	
CC	reactions. Compounds that modulate expression of the transcripts are	
CC	potentially useful for treatment of (metastatic) breast cancer, while	
CC	promoters from the transcripts are used to direct expression, in selected	
CC	cell types, of e.g. therapeutic genes (also ribozymes or antisense	
CC	sequences), particularly an antigen encoding sequence for use in gene or	
CC	cell-based vaccines. Polypeptides encoded by the transcripts are also	
CC	useful in vaccines; for diagnosing breast cancer and for raising	
CC	specific antibodies (Ab). Ab are used to detect the polypeptides or as	
CC	therapeutic agents. Host cells that produce the polypeptides can be used	
CC	to expand and isolate populations of educated, antigen-specific immune	
CC	effector cells, e.g. cytotoxic T lymphocytes, and these used for	
CC	adoptive immunotherapy.	
CC		
SQ	Sequence 10 BP; 3 A; 0 C; 1 G; 6 T; 0 other:	
OY		
Dd		
Query Match	100.0%; Score 7; DB 21; Length 10;	
Best Local Similarity	100.0%; Pred. No. 4.3e+04;	
Matches 7; Conservative	0; Mismatches 0; Indels 0; Gaps 0.	
OY	1 ttgata 7	
Dd	3 ttgtga 9	
RESULT		
AAZ82567		
ID	AAZ82567 standard; DNA; 10 BP.	
AC	AAZ82567;	
DT	07-APR-2000 (first entry)	
XX		
DE	Metastatic breast tumour cell upregulated transcript tag #1801.	

```

XX XX Human: metastatic breast tumour tissue; breast cancer; tag; primer;
KM KM non-metastatic breast tumour tissue; gene therapy; anticancer;
KW KW antimetastatic; vaccine; diagnosis; ss.
XX XX
OS OS Homo sapiens.
XX XX
PM PM W09965928-A2.
XX XX
PD PD 23-DEC-1999.
XX XX
PF PF 18-JUN-1999; 99WO-US13647.
XX XX
PR PR 19-JUN-1998; 98US-0089853.
PR PR 19-JUN-1998; 98US-0089897.
PR PR 19-JUN-1998; 98US-0090039.
PR PR 19-JUN-1998; 98US-0090040.
PR PR 19-JUN-1998; 98US-0090041.
XX XX
PA PA (GENZ ) GENZYME CORP.
PA PA (ROBE) ROBERTS B L.
PA PA (SHANV) SHANKARA S.
PI PI Roberts B., Shankara S;
XX XX
DR DR WPI: 2000-106079/09.
XX XX
PT PT Isolated polynucleotides differentially expressed between metastatic
PT PT and non-metastatic breast cancer cells, useful for diagnosis,
XX XX prevention and treatment of cancer -
XX XX
PS PS Claim 1; Page 107; 21pp; English.
XX XX
CC CC AA280767 to AA283941 represent tags corresponding to distinct
CC CC transcripts that are preferentially transcribed in the metastatic breast
CC CC tumour tissue (i.e. are upregulated in metastatic breast tumour cells).
CC CC AA283942 to AA286677 represent tags corresponding to distinct transcripts
CC CC that are preferentially transcribed in the primary or non-metastatic
CC CC breast tumour tissue (i.e. are downregulated in metastatic breast tumour
CC CC cells). These transcripts can be used for diagnosis, prognosis,
CC CC monitoring and treatment of breast cancer, particularly where metastatic.
CC CC diagnosis is by standard immunoassays or hybridisation/amplification
CC CC reactions. Compounds that modulate expression of the transcripts are
CC CC potentially useful for treatment of (metastatic) breast cancer, while
CC CC promoters from the transcripts are used to direct expression, in selected
CC CC cell types, of e.g. therapeutic genes (also ribozymes or antisense
CC CC sequences), particularly an antigen-encoding sequence for use in gene or
CC CC cell-based vaccines. Polypeptides encoded by the transcripts are also
CC CC useful in vaccines; for diagnosing breast cancer and for raising
CC CC specific antibodies (Ab). Ab are used to detect the polypeptides or as
CC CC therapeutic agents. Host cells that produce the polypeptides can be used
CC CC to expand and isolate populations of educated, antigen-specific immune
CC CC effector cells, e.g. cytotoxic T lymphocytes, and these used for
CC CC adoptive immunotherapy.
XX XX
SQ SQ Sequence 10 BP; 3 A; 0 C; 2 G; 5 T; 0 other;

Query Match 100.0%; Score 7; DB 21; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.3e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ttgata 7
| | | | | | |
Db 1 ttgata 7

RESULT 5
ID AA282768 standard; DNA; 10 BP.
XX AC AA282768;
XX XX

```

DT 07-APR-2000 (first entry)  
XX Metastatic breast tumour cell upregulated transcript tag #2002.  
DE  
XX  
XX Human: metastatic breast tumour tissue; breast cancer; tag; primer;  
KM non-metastatic breast tumour tissue; gene therapy; anticancer;  
KW antimetastatic; vaccine; diagnosis; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO965928-A2.  
XX  
PD 23-DEC-1999.  
XX  
PF 18-JUN-1999; 99MO-US13647.  
XX  
PR 19-JUN-1998; 98US-0089853.  
XX 19-JUN-1998; 98US-0089997.  
PR 19-JUN-1998; 98US-0090039.  
PR 19-JUN-1998; 98US-0090040.  
PR 19-JUN-1998; 98US-0090041.  
XX  
PA (GENZ ) GENZYME CORP.  
PA (ROBE/) ROBERTS B L.  
PA (SHAN/) SHANKARA S.  
XX  
PI Roberts BL, Shankara S;  
XX  
DR WPI: 2000-106079/09.  
XX  
PT Isolated polynucleotides differentially expressed between metastatic  
PT and non-metastatic breast cancer cells, useful for diagnosis,  
PT prevention and treatment of cancer -  
XX  
PS Claim 1; Page 113; 219pp; English.  
XX  
CC AA280767 to AA283941 represent tags corresponding to distinct  
CC transcripts that are preferentially transcribed in the metastatic breast  
CC tumour tissue (i.e. are upregulated in metastatic breast tumour cells).  
CC AA283942 to AA286677 represent tags corresponding to distinct transcripts  
CC that are preferentially transcribed in the primary or non-metastatic  
CC breast tumour tissue (i.e. are downregulated in metastatic breast tumour  
CC cells). These transcripts can be used for diagnosis, prognosis,  
CC monitoring and treatment of breast cancer, particularly where metastatic.  
CC reactions. Compounds that modulate expression of the transcripts are  
CC potentially useful for treatment of (metastatic) breast cancer, while  
CC promoters from the transcripts are used to direct expression, in selected  
CC cell types, of e.g. therapeutic genes (also ribozymes or antisense  
CC sequences), particularly an antigen-encoding sequence for use in gene or  
CC cell-based vaccines. Polypeptides encoded by the transcripts are also  
CC useful in vaccines for diagnosing breast cancer and for raising  
CC specific antibodies (Ab). Ab are used to detect the polypeptides or as  
CC therapeutic agents. Host cells that produce the polypeptides can be used  
CC to expand and isolate populations of educated, antigen-specific immune  
CC effector cells, e.g. cytotoxic T lymphocytes, and these used for  
CC adoptive immunotherapy.  
XX  
SQ Sequence 10 BP; 4 A; 0 C; 2 G; 4 T; 0 other;  
  
Query Match 100.0%; Score 7; DB 21; Length 10;  
Best Local Similarity 100.0%; Pred. No. 4.3e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 ttgtata 7  
| | | | | | |  
Db 2 ttgtata 8

RESULT 6  
AA283117  
ID AA283117 standard; DNA: 10 BP.

XX  
XX AA283117;  
AC  
XX  
DT 07-APR-2000 (first entry)  
XX  
XX Metastatic breast tumour cell upregulated transcript tag #2351.  
DE  
XX  
XX Human: metastatic breast tumour tissue; breast cancer; tag; primer;  
KM non-metastatic breast tumour tissue; gene therapy; anticancer;  
KW antimetastatic; vaccine; diagnosis; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO965928-A2.  
XX  
PD 23-DEC-1999.  
XX  
PF 18-JUN-1999; 99MO-US13647.  
XX  
PR 19-JUN-1998; 98US-0089853.  
XX 19-JUN-1998; 98US-0089997.  
PR 19-JUN-1998; 98US-0090039.  
PR 19-JUN-1998; 98US-0090040.  
PR 19-JUN-1998; 98US-0090041.  
XX  
PA (GENZ ) GENZYME CORP.  
PA (ROBE/) ROBERTS B L.  
PA (SHAN/) SHANKARA S.  
XX  
PI Roberts BL, Shankara S;  
XX  
DR WPI: 2000-106079/09.  
XX  
PT Isolated polynucleotides differentially expressed between metastatic  
PT and non-metastatic breast cancer cells, useful for diagnosis,  
PT prevention and treatment of cancer -  
XX  
PS Claim 1; Page 122; 219pp; English.  
XX  
CC AA280767 to AA283941 represent tags corresponding to distinct  
CC transcripts that are preferentially transcribed in the metastatic breast  
CC tumour tissue (i.e. are upregulated in metastatic breast tumour cells).  
CC AA283942 to AA286677 represent tags corresponding to distinct transcripts  
CC that are preferentially transcribed in the primary or non-metastatic  
CC breast tumour tissue (i.e. are downregulated in metastatic breast tumour  
CC cells). These transcripts can be used for diagnosis, prognosis,  
CC monitoring and treatment of breast cancer, particularly where metastatic.  
CC reactions. Compounds that modulate expression of the transcripts are  
CC potentially useful for treatment of (metastatic) breast cancer, while  
CC promoters from the transcripts are used to direct expression, in selected  
CC cell types, of e.g. therapeutic genes (also ribozymes or antisense  
CC sequences), particularly an antigen-encoding sequence for use in gene or  
CC cell-based vaccines. Polypeptides encoded by the transcripts are also  
CC useful in vaccines for diagnosing breast cancer and for raising  
CC specific antibodies (Ab). Ab are used to detect the polypeptides or as  
CC therapeutic agents. Host cells that produce the polypeptides can be used  
CC to expand and isolate populations of educated, antigen-specific immune  
CC effector cells, e.g. cytotoxic T lymphocytes, and these used for  
CC adoptive immunotherapy.  
XX  
SQ Sequence 10 BP; 4 A; 0 C; 1 G; 5 T; 0 other;

Query Match 100.0%; Score 7; DB 21; Length 10;  
Best Local Similarity 100.0%; Pred. No. 4.3e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 ttgtata 7  
| | | | | | |  
Db 1 ttgtata 7

RESULT 7  
AAF34548/C  
ID AAF34548 standard; DNA: 10 BP.  
XX  
AC AAF34548;  
XX  
DT 23-MAR-2001 (first entry)  
XX  
DE Yeast NORF gene SAGE tag oligonucleotide seq ID NO:1287.  
XX  
KM Yeast: Saccharomyces cerevisiae; characterisation: cell cycle; NORF;  
KM nor previously assigned open reading frame; nonannotated ORF; SAGE;  
KM Serial analysis of gene expression; antifungal; tag; identification;  
KM linker; PCR primer; ds..  
XX  
OS Saccharomyces cerevisiae.  
XX  
PN W0200077214-A2.  
XX  
PD 21-DEC-2000.  
XX  
PF 14-JUN-2000; 2000WO-US16223.  
XX  
PR 16-JUN-1999; 99US-0335032.  
XX  
PA (UYJO ) UNIV JOHNS HOPKINS.  
XX  
PI Velculescu V, Vogelstein B, Kinzler K;  
XX WPI: 2001-061874/07.  
XX  
PT Yeast gene coding sequences comprising NORF genes with serial analysis  
PT of gene expression (SAGE) tags, useful for studying, monitoring and  
PT affecting phases of the cell cycle -  
XX  
PS Example; Page 46; 419pp; English.  
XX  
CC The present invention describes an isolated DNA molecule comprising a  
CC coding sequence of a yeast gene selected from a group of 745 NORF (not  
CC previously assigned open reading frame; or nonannotated ORF) genes  
CC comprising a SAGE (serial analysis of gene expression) tag. Also  
CC described are: (1) a method (M1) of using NORF genes to affect the cell  
CC cycle comprising administering a NORF gene whose expression varies by at  
CC least 10% between any two phases of the cell cycle selected from log  
CC phase, S phase and G2/M; (2) a method (M2) for screening candidate  
CC antifungal drugs comprising: (a) contacting a test substance with a  
CC yeast cell; and (b) monitoring expression of a NORF gene whose  
CC expression varies as in M1, where a test substance which modifies the  
CC expression of the yeast gene is a candidate antifungal drug; (3) a method  
CC (M3) for identifying human genes which are involved in cell cycle  
CC progression comprising contacting human DNA with a probe which comprises  
CC at least 10 contiguous nucleotides of a NORF gene whose expression varies  
CC as in M1; and (4) a method (M4) for identifying a candidate drug as a  
CC member of a class of drugs having a characteristic effect on gene  
CC expression in a yeast cell comprising contacting a yeast cell with a  
CC candidate drug and monitoring expression in the yeast cell of at least 1  
CC NORF gene whose expression is affected by the class of drugs. The NORF  
CC genes may be used to study, monitor and affect phases of the cell cycle,  
CC the differentially expressed genes may be used as markers of phases of  
CC the cell cycle. The methods may be used to identify candidate drugs which  
CC affect the cell cycle and for identification of antifungal drugs.  
CC AAF33268 to AAF44064 represent SAGE tags used in the exemplification of  
CC the present invention. AAF33262 to AAF33267 represent linkers and PCR  
CC primers used in the SAGE method, in the exemplification of the present  
CC invention.  
XX  
SQ Sequence 10 BP; 7 A; 1 C; 0 G; 2 T; 0 other;

Qy 1 ttctgata 7  
Db 8 TTTGATA 2  
RESULT 8  
AAF35054/C  
ID AAF35054 standard; DNA: 10 BP.  
XX  
AC AAF35054;  
XX  
DT 23-MAR-2001 (first entry)  
XX  
DE Yeast NORF gene SAGE tag oligonucleotide seq ID NO:1793.  
XX  
KM Yeast: Saccharomyces cerevisiae; characterisation: cell cycle; NORF;  
KM nor previously assigned open reading frame; nonannotated ORF; SAGE;  
KM Serial analysis of gene expression; antifungal; tag; identification;  
KM linker; PCR primer; ds..  
XX  
OS Saccharomyces cerevisiae.  
XX  
PN W0200077214-A2.  
XX  
PD 21-DEC-2000.  
XX  
PF 14-JUN-2000; 2000WO-US16223.  
XX  
PR 16-JUN-1999; 99US-0335032.  
XX  
PA (UYJO ) UNIV JOHNS HOPKINS.  
XX  
PI Velculescu V, Vogelstein B, Kinzler K;  
XX WPI: 2001-061874/07.  
XX  
PT Yeast gene coding sequences comprising NORF genes with serial analysis  
PT of gene expression (SAGE) tags, useful for studying, monitoring and  
PT affecting phases of the cell cycle -  
XX  
PS Example; Page 64; 419pp; English.  
XX  
CC The present invention describes an isolated DNA molecule comprising a  
CC coding sequence of a yeast gene selected from a group of 745 NORF (not  
CC previously assigned open reading frame; or nonannotated ORF) genes  
CC comprising a SAGE (serial analysis of gene expression) tag. Also  
CC described are: (1) a method (M1) of using NORF genes to affect the cell  
CC cycle comprising administering a NORF gene whose expression varies by at  
CC least 10% between any two phases of the cell cycle selected from log  
CC phase, S phase and G2/M; (2) a method (M2) for screening candidate  
CC antifungal drugs comprising: (a) contacting a test substance with a  
CC yeast cell; and (b) monitoring expression of a NORF gene whose  
CC expression varies as in M1, where a test substance which modifies the  
CC expression of the yeast gene is a candidate antifungal drug; (3) a method  
CC (M3) for identifying human genes which are involved in cell cycle  
CC progression comprising contacting human DNA with a probe which comprises  
CC at least 10 contiguous nucleotides of a NORF gene whose expression varies  
CC as in M1; and (4) a method (M4) for identifying a candidate drug as a  
CC member of a class of drugs having a characteristic effect on gene  
CC expression in a yeast cell comprising contacting a yeast cell with a  
CC candidate drug and monitoring expression in the yeast cell of at least 1  
CC NORF gene whose expression is affected by the class of drugs. The NORF  
CC genes may be used to study, monitor and affect phases of the cell cycle,  
CC the differentially expressed genes may be used as markers of phases of  
CC the cell cycle. The methods may be used to identify candidate drugs which  
CC affect the cell cycle and for identification of antifungal drugs.  
CC AAF33268 to AAF44064 represent SAGE tags used in the exemplification of  
CC the present invention. AAF33262 to AAF33267 represent linkers and PCR  
CC primers used in the SAGE method, in the exemplification of the present  
CC invention.  
XX  
SQ Sequence 10 BP; 6 A; 2 C; 0 G; 2 T; 0 other;

Query Match 100.0%; Score 7; DB 22; Length 10;  
Best Local Similarity 100.0%; Pred. No. 4.3e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgtgata 7  
| | | | |  
DB 8 TTTGATA 2

RESULT 9  
AAF38092/c  
ID AAF38092 standard; DNA; 10 BP.

XX AAF38092;

DT 23-MAR-2001 (first entry)

XX Yeast NORF gene SAGE tag oligonucleotide SEQ ID NO:4831.

XX Yeast: Saccharomyces cerevisiae; characterisation: cell cycle: NORF;  
KW nor previously assigned open reading frame; nonannotated ORF; SAGE;  
KW serial analysis of gene expression; antifungal; tag; identification;  
KW linker; PCR primer; ds.

XX Saccharomyces cerevisiae.

XX WO20007214-A2.

XX 21-DEC-2000.

XX 14-JUN-2000; 2000WO-US16223.

XX 16-JUN-1999; 99US-0335032.

XX (UYJO ) UNIV JOHNS HOPKINS.

XX Velculescu V, Vogelstein B, Kinzler K;

XX WPI: 2001-061874/07.

PT Yeast gene coding sequences comprising NORF genes with serial analysis  
PT of gene expression (SAGE) tags, useful for studying, monitoring and  
PT affecting phases of the cell cycle -

XX Example; Page 172; 419pp; English.

XX The present invention describes an isolated DNA molecule comprising a  
XX coding sequence of a yeast gene selected from a group of 745 NORF (not  
XX previously assigned open reading frame; or nonannotated ORF) genes  
XX comprising a SAGE (serial analysis of gene expression) tag. Also  
XX described are: (1) a method (M1) of using NORF genes to affect the cell  
XX cycle comprising administering a NORF gene whose expression varies by at  
XX least 10% between any two phases of the cell cycle selected from log  
XX phase, S phase and G2/M; (2) a method (M2) for screening candidate  
XX antifungal drugs comprising: (a) contacting a test substance with a  
XX yeast cell; and (b) monitoring expression of a NORF gene whose  
XX expression varies as in M1, where a test substance which modifies the  
XX expression of the yeast gene is a candidate antifungal drug; (3) a method  
XX (M3) for identifying human genes which are involved in cell cycle  
XX progression comprising contacting human DNA with a probe which comprises  
XX at least 10 contiguous nucleotides of a NORF gene whose expression varies  
XX as in M1; and (4) a method (M4) for identifying a candidate drug as a  
XX member of a class of drugs having a characteristic effect on gene  
XX expression in a yeast cell comprising contacting a yeast cell with a  
XX candidate drug and monitoring expression in the yeast cell of at least 1  
XX NORF gene whose expression is affected by the class of drugs. The NORF  
XX genes may be used to study, monitor and affect phases of the cell cycle,  
XX the differentially expressed genes may be used as markers of phases of  
XX the cell cycle. The methods may be used to identify candidate drugs which  
XX affect the cell cycle and for identification of antifungal drugs.  
XX AAF33268 to AAF44064 represent SAGE tags used in the exemplification of  
XX the present invention. AAF33262 to AAF33267 represent linkers and PCR

CC primers used in the SAGE method, in the exemplification of the present  
CC invention.

SQ Sequence 10 BP; 4 A; 2 C; 0 G; 4 T; 0 other;

Query Match 100.0%; Score 7; DB 22; Length 10;  
Best Local Similarity 100.0%; Pred. No. 4.3e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgtgata 7  
| | | | |  
DB 7 TTTGATA 1

RESULT 10  
AAF39829/c  
ID AAF39829 standard; DNA; 10 BP.

XX AAF39829;

DT 23-MAR-2001 (first entry)

XX Yeast NORF gene SAGE tag oligonucleotide SEQ ID NO:6568.

XX Yeast: Saccharomyces cerevisiae; characterisation: cell cycle: NORF;  
KW nor previously assigned open reading frame; nonannotated ORF; SAGE;  
KW serial analysis of gene expression; antifungal; tag; identification;  
KW linker; PCR primer; ds.

XX Saccharomyces cerevisiae.

XX WO20007214-A2.

XX 21-DEC-2000.

XX 14-JUN-2000; 2000WO-US16223.

XX 16-JUN-1999; 99US-0335032.

XX (UYJO ) UNIV JOHNS HOPKINS.

XX Velculescu V, Vogelstein B, Kinzler K;

XX WPI: 2001-061874/07.

PT Yeast gene coding sequences comprising NORF genes with serial analysis  
PT of gene expression (SAGE) tags, useful for studying, monitoring and  
PT affecting phases of the cell cycle -

XX Example; Page 234; 419pp; English.

XX The present invention describes an isolated DNA molecule comprising a  
XX coding sequence of a yeast gene selected from a group of 745 NORF (not  
XX previously assigned open reading frame; or nonannotated ORF) genes  
XX comprising a SAGE (serial analysis of gene expression) tag. Also  
XX described are: (1) a method (M1) of using NORF genes to affect the cell  
XX cycle comprising administering a NORF gene whose expression varies by at  
XX least 10% between any two phases of the cell cycle selected from log  
XX phase, S phase and G2/M; (2) a method (M2) for screening candidate  
XX antifungal drugs comprising: (a) contacting a test substance with a  
XX yeast cell; and (b) monitoring expression of a NORF gene whose  
XX expression varies as in M1, where a test substance which modifies the  
XX expression of the yeast gene is a candidate antifungal drug; (3) a method  
XX (M3) for identifying human genes which are involved in cell cycle  
XX progression comprising contacting human DNA with a probe which comprises  
XX at least 10 contiguous nucleotides of a NORF gene whose expression varies  
XX as in M1; and (4) a method (M4) for identifying a candidate drug as a  
XX member of a class of drugs having a characteristic effect on gene  
XX expression in a yeast cell comprising contacting a yeast cell with a  
XX candidate drug and monitoring expression in the yeast cell of at least 1  
XX NORF gene whose expression is affected by the class of drugs. The NORF  
XX genes may be used to study, monitor and affect phases of the cell cycle,

CC the differentially expressed genes may be used as markers of phases of  
CC the cell cycle. The methods may be used to identify candidate drugs which  
CC affect the cell cycle and for identification of antifungal drugs.  
CC AAF33268 to AAF44064 represent SAGE tags used in the exemplification of  
CC the present invention. AAF33262 to AAF33267 represent linkers and PCR  
CC primers used in the SAGE method, in the exemplification of the present  
CC invention.  
XX  
XX Sequence 10 BP; 5 A; 1 C; 1 G; 3 T; 0 other;  
SO  
Query Match 100.0%; Score 7; DB 22; Length 10;  
Best Local Similarity 100.0%; Pred. No. 4.3e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 ttgtata 7  
| | | | |  
Db 10 TTTGATA 4  
RESULT 11  
AAF40369/c  
ID AAF40369 standard; DNA: 10 BP.  
XX  
XX AAF40369;  
AC  
DT 23-MAR-2001 (first entry)  
XX  
DE Yeast NORF gene SAGE tag oligonucleotide SEQ ID NO:7108.  
XX  
XX Yeast; Saccharomyces cerevisiae; characterisation: cell cycle; NORF;  
KM nor previously assigned open reading frame; nonannotated ORF; SAGE;  
KW serial analysis of gene expression; antifungal; tag; identification;  
KW linker: PCR primer; ds.  
XX  
XX Saccharomyces cerevisiae.  
OS  
PN WO200077214-A2.  
XX  
PD 21-DEC-2000.  
XX  
PF 14-JUN-2000; 2000MO-US16223.  
XX  
XX 16-JUN-1999; 99US-0335032.  
PR  
PA (UYJO ) UNIV JOHNS HOPKINS.  
XX  
PI Velculescu V, Vogelstein B, Kinzler K;  
XX  
DR WPI; 2001-061874/07.  
XX  
XX  
PT Yeast gene coding sequences comprising NORF genes with serial analysis  
PT of gene expression (SAGE) tags, useful for studying, monitoring and  
PT affecting phases of the cell cycle -  
XX  
XX Example; Page 253; 419pp; English.  
PS  
XX The present invention describes an isolated DNA molecule comprising a  
CC coding sequence of a yeast gene selected from a group of 745 NORF (not  
CC previously assigned open reading frame; or nonannotated ORF) genes  
CC comprising a SAGE (serial analysis of gene expression) tag. Also  
CC described are: (1) a method (M1) of using NORF genes to affect the cell  
CC cycle comprising administering a NORF gene whose expression varies by at  
CC least 10% between any two phases of the cell cycle selected from log  
CC phase, S phase and G2/M; (2) a method (M2) for screening candidate  
CC antifungal drugs comprising: (a) contacting a test substance with a  
CC yeast cell; and (b) monitoring expression of a NORF gene whose  
CC expression varies as in M1, where a test substance which modifies the  
CC expression of the yeast gene is a candidate antifungal drug; (3) a method  
CC (M3) for identifying human genes which are involved in cell cycle  
CC progression comprising contacting human DNA with a probe which comprises  
CC at least 10 contiguous nucleotides of a NORF gene whose expression varies  
CC as in M1; and (4) a method (M4) for identifying a candidate drug as a

CC member of a class of drugs having a characteristic effect on gene  
CC expression in a yeast cell comprising contacting a yeast cell with a  
CC candidate drug and monitoring expression in the yeast cell of at least 1  
CC NORF gene whose expression is affected by the class of drugs. The NORF  
CC genes may be used to study, monitor and affect phases of the cell cycle,  
CC the differentially expressed genes may be used as markers of phases of  
CC the cell cycle. The methods may be used to identify candidate drugs which  
CC affect the cell cycle and for identification of antifungal drugs.  
CC AAF33268 to AAF44064 represent SAGE tags used in the exemplification of  
CC the present invention. AAF33262 to AAF33267 represent linkers and PCR  
CC primers used in the SAGE method, in the exemplification of the present  
CC invention.  
XX  
XX Sequence 10 BP; 5 A; 1 C; 0 G; 4 T; 0 other;  
SO  
Query Match 100.0%; Score 7; DB 22; Length 10;  
Best Local Similarity 100.0%; Pred. No. 4.3e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 ttgtata 7  
| | | | |  
Db 10 TTTGATA 4  
RESULT 12  
AAF40389  
ID AAF40389 standard; DNA: 10 BP.  
XX  
XX AAF40389;  
AC  
DT 23-MAR-2001 (first entry)  
XX  
DE Yeast NORF gene SAGE tag oligonucleotide SEQ ID NO:7128.  
XX  
XX Yeast; Saccharomyces cerevisiae; characterisation: cell cycle; NORF;  
KM nor previously assigned open reading frame; nonannotated ORF; SAGE;  
KW serial analysis of gene expression; antifungal; tag; identification;  
KW linker: PCR primer; ds.  
XX  
XX Saccharomyces cerevisiae.  
OS  
PN WO200077214-A2.  
XX  
PD 21-DEC-2000.  
XX  
PF 14-JUN-2000; 2000MO-US16223.  
XX  
XX 16-JUN-1999; 99US-0335032.  
PR  
PA (UYJO ) UNIV JOHNS HOPKINS.  
XX  
PI Velculescu V, Vogelstein B, Kinzler K;  
XX  
DR WPI; 2001-061874/07.  
XX  
XX  
PT Yeast gene coding sequences comprising NORF genes with serial analysis  
PT of gene expression (SAGE) tags, useful for studying, monitoring and  
PT affecting phases of the cell cycle -  
XX  
XX Example; Page 254; 419pp; English.  
PS  
XX The present invention describes an isolated DNA molecule comprising a  
CC coding sequence of a yeast gene selected from a group of 745 NORF (not  
CC previously assigned open reading frame; or nonannotated ORF) genes  
CC comprising a SAGE (serial analysis of gene expression) tag. Also  
CC described are: (1) a method (M1) of using NORF genes to affect the cell  
CC cycle comprising administering a NORF gene whose expression varies by at  
CC least 10% between any two phases of the cell cycle selected from log  
CC phase, S phase and G2/M; (2) a method (M2) for screening candidate  
CC antifungal drugs comprising: (a) contacting a test substance with a  
CC yeast cell; and (b) monitoring expression of a NORF gene whose  
CC expression varies as in M1, where a test substance which modifies the

expression of the yeast gene is a candidate antifungal drug; (3) a method (M3) for identifying human genes which are involved in cell cycle progression comprising contacting human DNA with a probe which comprises at least 10 contiguous nucleotides of a NORF gene whose expression varies as in M1; and (4) a method (M4) for identifying a candidate drug as a member of a class of drugs having a characteristic effect on gene expression in a yeast cell comprising contacting a yeast cell with a candidate drug and monitoring expression in the yeast cell of at least 1 NORF gene whose expression is affected by the class of drugs. The NORF genes may be used to study, monitor and affect phases of the cell cycle, the differentially expressed genes may be used as markers of phases of the cell cycle. The methods may be used to identify candidate drugs which affect the cell cycle and for identification of antifungal drugs.

AAF33268 to AAF44064 represent SAGE tags used in the exemplification of the present invention. AAF33262 to AAF33267 represent linkers and PCR primers used in the SAGE method, in the exemplification of the present invention.

Sequence 10 BP; 3 A; 1 C; 1 G; 5 T; 0 other;

Query Match 100.0%; Score 7; DB 22; Length 10;  
Best Local Similarity 100.0%; Pred. No. 4.3e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ttgtgata 7  
|||||||  
Db 2 ttgtgata 8

## RESULT 13

AAF40414  
ID AAF40414 standard; DNA; 10 BP.

AC AAF40414;

DT 23-MAR-2001 (first entry)

DE Yeast NORF gene SAGE tag oligonucleotide SEQ ID NO:7153.

XX Yeast: Saccharomyces cerevisiae; characterisation; cell cycle; NORF;  
KW nor previously assigned open reading frame; nonannotated ORF; SAGE;  
KW serial analysis of gene expression; antifungal; tag; identification;  
KW linker; PCR primer; ds.

OS Saccharomyces cerevisiae.

PN WO200077214-A2.

PD 21-DEC-2000.

PF 14-JUN-2000; 2000WO-US16223.

PR 16-JUN-1999; 99US-0335032.

PA (UYJO ) UNIV JOHNS HOPKINS.

PI Velculescu V, Vogelstein B, Kinzler K;

DR WPI: 2001-061874/07.

XX Yeast gene coding sequences comprising NORF genes with serial analysis  
PT of gene expression (SAGE) tags, useful for studying, monitoring and  
PT affecting phases of the cell cycle -

PS Example: Page 255; 419pp; English.

XX The present invention describes an isolated DNA molecule comprising a  
CC coding sequence of a yeast gene selected from a group of 745 NORF (not  
CC previously assigned open reading frame; or nonannotated ORF) genes  
CC comprising a SAGE (serial analysis of gene expression) tag. Also  
CC described are: (1) a method (M1) of using NORF genes to affect the cell  
cycle comprising administering a NORF gene whose expression varies by at

least 10% between any two phases of the cell cycle selected from log  
CC phase, S phase and G2/M; (2) a method (M2) for screening candidate  
CC antifungal drugs comprising: (a) contacting a test substance with a  
CC yeast cell; and (b) monitoring expression of a NORF gene whose  
CC expression varies as in M1, where a test substance which modifies the  
CC expression of the yeast gene is a candidate antifungal drug; (3) a method  
CC (M3) for identifying human genes which are involved in cell cycle  
CC progression comprising contacting human DNA with a probe which comprises  
CC at least 10 contiguous nucleotides of a NORF gene whose expression varies  
CC as in M1; and (4) a method (M4) for identifying a candidate drug as a  
CC member of a class of drugs having a characteristic effect on gene  
CC expression in a yeast cell comprising contacting a yeast cell with a  
CC candidate drug and monitoring expression in the yeast cell of at least 1  
CC NORF gene whose expression is affected by the class of drugs. The NORF  
CC genes may be used to study, monitor and affect phases of the cell cycle,  
CC the differentially expressed genes may be used as markers of phases of  
CC the cell cycle. The methods may be used to identify candidate drugs which  
CC affect the cell cycle and for identification of antifungal drugs.

AAF33268 to AAF44064 represent SAGE tags used in the exemplification of the present invention. AAF33262 to AAF33267 represent linkers and PCR primers used in the SAGE method, in the exemplification of the present invention.

Sequence 10 BP; 3 A; 1 C; 1 G; 5 T; 0 other;

Query Match 100.0%; Score 7; DB 22; Length 10;  
Best Local Similarity 100.0%; Pred. No. 4.3e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ttgtgata 7  
|||||||  
Db 3 ttgtgata 9

## RESULT 14

AAF40778  
ID AAF40778 standard; DNA; 10 BP.

AC AAF40778;

DT 23-MAR-2001 (first entry)

DE Yeast NORF gene SAGE tag oligonucleotide SEQ ID NO:7517.

XX Yeast: Saccharomyces cerevisiae; characterisation; cell cycle; NORF;  
KW nor previously assigned open reading frame; nonannotated ORF; SAGE;  
KW serial analysis of gene expression; antifungal; tag; identification;  
KW linker; PCR primer; ds.

OS Saccharomyces cerevisiae.

PN WO200077214-A2.

PD 21-DEC-2000.

PF 14-JUN-2000; 2000WO-US16223.

PR 16-JUN-1999; 99US-0335032.

PA (UYJO ) UNIV JOHNS HOPKINS.

PI Velculescu V, Vogelstein B, Kinzler K;

DR WPI: 2001-061874/07.

XX Yeast gene coding sequences comprising NORF genes with serial analysis  
PT of gene expression (SAGE) tags, useful for studying, monitoring and  
PT affecting phases of the cell cycle -

PS Example: Page 268; 419pp; English.

XX The present invention describes an isolated DNA molecule comprising a

CC coding sequence of a yeast gene selected from a group of 745 NORF (not  
 CC previously assigned open reading frame; or nonannotated ORF) genes  
 CC comprising a SAGE (serial analysis of gene expression) tag. Also  
 CC described are: (1) a method (M1) of using NORF genes to affect the cell  
 CC cycle comprising administering a NORF gene whose expression varies by at  
 CC least 10% between any two phases of the cell cycle selected from log  
 CC phase, S phase and G2/M; (2) a method (M2) for screening candidate  
 CC antifungal drugs comprising: (a) contacting a test substance with a  
 CC yeast cell; and (b) monitoring expression of a NORF gene whose  
 CC expression varies as in M1, where a test substance which modifies the  
 CC expression of the yeast gene is a candidate antifungal drug; (3) a method  
 CC (M3) for identifying human genes which are involved in cell cycle  
 CC progression comprising contacting human DNA with a probe which comprises  
 CC at least 10 contiguous nucleotides of a NORF gene whose expression varies  
 CC as in M1; and (4) a method (M4) for identifying a candidate drug as a  
 CC member of a class of drugs having a characteristic effect on gene  
 CC expression in a yeast cell comprising contacting a yeast cell with a  
 CC candidate drug and monitoring expression in the yeast cell of at least 1  
 CC NORF gene whose expression is affected by the class of drugs. The NORF  
 CC genes may be used to study, monitor and affect phases of the cell cycle,  
 CC the differentially expressed genes may be used as markers of phases of  
 CC the cell cycle. The methods may be used to identify candidate drugs which  
 CC affect the cell cycle and for identification of antifungal drugs.  
 CC AAF33268 to AAF44064 represent SAGE tags used in the exemplification of  
 CC the present invention. AAF33262 to AAF33267 represent linkers and PCR  
 CC primers used in the SAGE method, in the exemplification of the present  
 CC invention.  
 CC XX  
 S0 Sequence 10 BP; 3 A; 1 C; 2 G; 4 T; 0 other;

Query Match 100.0%; Score 7; DB 22; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 4.3e+04;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgata 7  
 |||||  
 Db 3 ttgata 9

RESULT 15  
 AAF41077  
 ID AAF41077 standard; DNA; 10 BP.  
 XX  
 AC AAF41077;  
 XX  
 DT 23-MAR-2001 (first entry)  
 XX  
 DE Yeast NORF gene SAGE tag oligonucleotide SEQ ID NO:7816.  
 XX  
 KW Yeast; Saccharomyces cerevisiae; characterisation; cell cycle; NORF;  
 KW not previously assigned open reading frame; nonannotated ORF; SAGE;  
 KW serial analysis of gene expression; antifungal; tag; identification;  
 KW linker; PCR primer; ds.  
 XX  
 OS Saccharomyces cerevisiae.  
 XX  
 PN WO200077214-A2.  
 XX  
 PD 21-DEC-2000.  
 XX  
 PE 14-JUN-2000; 2000WO-US16223.  
 XX  
 PR 16-JUN-1999; 99US-0335032.  
 XX  
 PA (UYJO ) UNIV JOHNS HOPKINS.  
 PI Velculescu V, Vogelstein B, Kinzler K;  
 XX  
 DR WPI: 2001-061874/07.  
 XX  
 PT Yeast gene coding sequences comprising NORF genes with serial analysis  
 PT of gene expression (SAGE) tags, useful for studying, monitoring and

PT affecting phases of the cell cycle -  
 XX  
 PS Example; Page 279; 419pp; English.  
 XX

CC The present invention describes an isolated DNA molecule comprising a  
 CC coding sequence of a yeast gene selected from a group of 745 NORF (not  
 CC previously assigned open reading frame; or nonannotated ORF) genes  
 CC comprising a SAGE (serial analysis of gene expression) tag. Also  
 CC described are: (1) a method (M1) of using NORF genes to affect the cell  
 CC cycle comprising administering a NORF gene whose expression varies by at  
 CC least 10% between any two phases of the cell cycle selected from log  
 CC phase, S phase and G2/M; (2) a method (M2) for screening candidate  
 CC antifungal drugs comprising: (a) contacting a test substance with a  
 CC yeast cell; and (b) monitoring expression of a NORF gene whose  
 CC expression varies as in M1, where a test substance which modifies the  
 CC expression of the yeast gene is a candidate antifungal drug; (3) a method  
 CC (M3) for identifying human genes which are involved in cell cycle  
 CC progression comprising contacting human DNA with a probe which comprises  
 CC at least 10 contiguous nucleotides of a NORF gene whose expression varies  
 CC as in M1; and (4) a method (M4) for identifying a candidate drug as a  
 CC member of a class of drugs having a characteristic effect on gene  
 CC expression in a yeast cell comprising contacting a yeast cell with a  
 CC candidate drug and monitoring expression in the yeast cell of at least 1  
 CC NORF gene whose expression is affected by the class of drugs. The NORF  
 CC genes may be used to study, monitor and affect phases of the cell cycle,  
 CC the differentially expressed genes may be used as markers of phases of  
 CC the cell cycle. The methods may be used to identify candidate drugs which  
 CC affect the cell cycle and for identification of antifungal drugs.  
 CC AAF33268 to AAF44064 represent SAGE tags used in the exemplification of  
 CC the present invention. AAF33262 to AAF33267 represent linkers and PCR  
 CC primers used in the SAGE method, in the exemplification of the present  
 CC invention.  
 CC XX  
 S0 Sequence 10 BP; 2 A; 2 C; 1 G; 5 T; 0 other;

Query Match 100.0%; Score 7; DB 22; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 4.3e+04;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgata 7  
 |||||  
 Db 3 ttgata 9

Search completed: December 15, 2001, 03:29:44  
 Job time: 7884 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 03:22:47 ; Search time 172.39 Seconds  
(without alignments)  
9.196 Million cell updates/sec

Title: US-09-380-826a-5  
Perfect score: 7  
Sequence: 1 ttgtgata 7

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 11323899 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/2/1na/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/1na/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/1na/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/1na/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/1na/PCRTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/1na/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	7	100.0	11 1	US-08-086-634-6	Sequence 6, Appli
C 2	7	100.0	11 1	US-08-086-634-7	Sequence 7, Appli
C 3	7	100.0	12 4	US-09-042-353-4	Sequence 4, Appli
C 4	7	100.0	12 4	US-08-758-417A-272	Sequence 5, Appli
C 5	7	100.0	13 1	US-08-166-664-5	Sequence 217, App
C 6	7	100.0	13 1	US-08-271-880A-217	Sequence 217, App
C 7	7	100.0	13 3	US-08-910-408-217	Sequence 217, App
C 8	7	100.0	13 3	US-09-249-215-217	Sequence 217, App
C 9	7	100.0	14 1	US-08-271-880A-212	Sequence 212, App
C 10	7	100.0	14 2	US-08-910-408-212	Sequence 212, App
C 11	7	100.0	14 3	US-09-249-215-212	Sequence 212, App
C 12	7	100.0	15 1	US-08-373-124A-90	Sequence 90, Appli
C 13	7	100.0	15 1	US-08-291-932A-129	Sequence 129, App
C 14	7	100.0	15 1	US-08-291-932A-130	Sequence 130, App
C 15	7	100.0	15 1	US-08-291-932A-131	Sequence 131, App
C 16	7	100.0	15 1	US-08-334-847-8	Sequence 8, Appli
C 17	7	100.0	15 1	US-08-334-847-9	Sequence 9, Appli
C 18	7	100.0	15 1	US-08-334-847-10	Sequence 10, Appli
C 19	7	100.0	15 1	US-08-334-847-655	Sequence 655, App
C 20	7	100.0	15 1	US-08-334-847-656	Sequence 656, App
C 21	7	100.0	15 1	US-08-435-628-90	Sequence 90, Appli
C 22	7	100.0	15 2	US-08-585-684B-113	Sequence 113, App
C 23	7	100.0	15 2	US-08-585-684B-114	Sequence 114, App
C 24	7	100.0	15 2	US-08-585-684B-115	Sequence 115, App
C 25	7	100.0	15 2	US-08-585-684B-116	Sequence 116, App
C 26	7	100.0	15 2	US-08-585-684B-2126	Sequence 2126, Ap
C 27	7	100.0	15 2	US-08-585-684B-2127	Sequence 2127, Ap

C 28	7	100.0	15 2	US-08-585-684B-2128	Sequence 2128, Ap
C 29	7	100.0	15 2	US-08-585-684B-2129	Sequence 2129, Ap
C 30	7	100.0	15 2	US-08-585-684B-2319	Sequence 2319, Ap
C 31	7	100.0	15 2	US-08-585-684B-2320	Sequence 2320, Ap
C 32	7	100.0	15 2	US-08-585-684B-2321	Sequence 2321, Ap
C 33	7	100.0	15 2	US-08-585-684B-2321	Sequence 2321, Ap
C 34	7	100.0	15 4	US-09-038-073-113	Sequence 113, App
C 35	7	100.0	15 4	US-09-038-073-114	Sequence 114, App
C 36	7	100.0	15 4	US-09-038-073-115	Sequence 115, App
C 37	7	100.0	15 4	US-09-038-073-116	Sequence 116, App
C 38	7	100.0	15 4	US-09-038-073-2126	Sequence 2126, App
C 39	7	100.0	15 4	US-09-038-073-2127	Sequence 2127, App
C 40	7	100.0	15 4	US-09-038-073-2128	Sequence 2128, App
C 41	7	100.0	15 4	US-09-038-073-2129	Sequence 2129, App
C 42	7	100.0	15 4	US-09-038-073-2319	Sequence 2319, App
C 43	7	100.0	15 4	US-09-038-073-2320	Sequence 2320, App
C 44	7	100.0	15 4	US-09-038-073-2321	Sequence 2321, App
C 45	7	100.0	16 3	US-08-930-917A-8	Sequence 8, Appli

## ALIGNMENTS

RESULT 1  
US-08-086-634-6/c  
Sequence 6, Application US/08086634  
Patent No. 5516634  
GENERAL INFORMATION:  
APPLICANT: NEWMAN, Peter J.  
APPLICANT: SANTOSO, Sencot Saleh  
TITLE OF INVENTION: Molecular Basis of the  
TITLE OF INVENTION: Human Platelet Brb/Brb Alloantigen System and  
TITLE OF INVENTION: Applications thereof  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Foley & Lardner  
STREET: P. O. Box 1497  
CITY: Madison  
STATE: Wisconsin  
COUNTRY: USA  
ZIP: 53701-1497  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Wordperfect, Version 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/086,634  
FILING DATE: 30-June-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Scanlon, William J.  
REGISTRATION NUMBER: 30,136  
REFERENCE/DOCKET NUMBER: 30383/132  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (608) 258-4284  
TELEFAX: (608) 258-4258  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA to mRNA  
US-08-086-634-6

Query Match 100.0%; Score 7; DB 1; Length 11;

Best Local Similarity 100.0%; Pred. No. 9.1e+03; Mismatches 0; Indels 0; Gaps 0;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Query 1 ttgtgata 7  
Db 7 TTTGATA 1

RESULT 2  
US-08-086-634-7/C  
Sequence 7, Application US/08086634  
Patent No. 5516634  
GENERAL INFORMATION:  
APPLICANT: NEWMAN, Peter J.  
APPLICANT: SANTOSO, Senot Saleh  
TITLE OF INVENTION: Molecular Basis of the  
TITLE OF INVENTION: Human Platelet Bra/Brb Alloantigen System and  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: P. O. Box 1497  
CITY: Madison  
STATE: Wisconsin  
COUNTRY: USA  
ZIP: 53701-1497  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WordPerfect, Version 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/086,634  
FILING DATE: 30-June-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Scanlon, William J.  
REGISTRATION NUMBER: 30,136  
REFERENCE/DOCKET NUMBER: 30383/132  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (608) 258-4284  
TELEFAX: (608) 258-4258  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
US-08-086-634-7

Query Match 100.0%; Score 7; DB 1; Length 11;  
Best Local Similarity 100.0%; Pred. No. 9.1e+03;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ttgtata 7  
Db 7 ttgtata 1

RESULT 3  
US-09-042-353-4  
Sequence 4, Application US/09042353  
Patent No. 6255458  
GENERAL INFORMATION:  
APPLICANT: Lonberg, Nils  
APPLICANT: KAY, Robert M.  
TITLE OF INVENTION: Transgenic No. 6255458-Human Animals for  
TITLE OF INVENTION: Producing Heterologous Antibodies  
NUMBER OF SEQUENCES: 421  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/042,353  
FILING DATE: 13-MAR-1998  
CLASSIFICATION: 800  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 07/810,279  
FILING DATE: 17-DEC-1991  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 07/853,408  
FILING DATE: 18-MAR-1992  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 07/904,068  
FILING DATE: 23-JUN-1992  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 07/990,860  
FILING DATE: 16-DEC-1992  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/053,131  
FILING DATE: 26-APR-1993  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/096,762  
FILING DATE: 22-JUL-1993  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/155,301  
FILING DATE: 18-NOV-1993  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/161,739  
FILING DATE: 03-DEC-1993  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/165,699  
FILING DATE: 10-DEC-1993  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/209,741  
FILING DATE: 09-MAR-1994  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/352,322  
FILING DATE: 07-DEC-1994  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/544,404  
FILING DATE: 10-OCT-1995  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/728,463  
FILING DATE: 10-OCT-1996  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/758,417  
FILING DATE: 02-DEC-1996  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US96/16433  
FILING DATE: 10-OCT-1996  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US97/21803  
FILING DATE: 01-DEC-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph T.  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 014643-009040US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-09-042-353-4

Query Match 100.0%; Score 7; DB 4; Length 12;

Best Local Similarity 100.0%; Pred. No. 9.1e+03;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgtata 7  
|||||  
Db 4 TTTGATA 10

RESULT 4  
US-08-758-417A-272

; Sequence 272, Application US/08758417A  
; Patent No. 6300129

; GENERAL INFORMATION:

; APPLICANT: Lomborg, Nils

; TITLE OF INVENTION: Transgenic No. 6300129-Human Animals for

; Producing Heterologous Antibodies

; NUMBER OF SEQUENCES: 417

; CORRESPONDENCE ADDRESSES:

; ADDRESSEE: Townsend and Townsend and Crew LLP

; STREET: Two Embarcadero Center, Eighth Floor

; CITY: San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94111-3834

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/758,417A

; FILING DATE: 02-Dec-1996

; CLASSIFICATION: <unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/728,463

; FILING DATE: 10-OCT-1996

; APPLICATION NUMBER: US 08/544,404

; FILING DATE: 10-OCT-1995

; APPLICATION NUMBER: US 08/352,322

; FILING DATE: 07-DEC-1994

; APPLICATION NUMBER: US 08/209,741

; FILING DATE: 09-MAR-1994

; APPLICATION NUMBER: US 08/165,699

; FILING DATE: 10-DEC-1993

; APPLICATION NUMBER: US 08/161,739

; FILING DATE: 03-DEC-1993

; APPLICATION NUMBER: US 08/155,301

; FILING DATE: 18-NOV-1993

; APPLICATION NUMBER: US 08/096,762

; FILING DATE: 22-JUL-1993

; APPLICATION NUMBER: US 08/053,131

; FILING DATE: 26-APR-1993

; APPLICATION NUMBER: US 07/990,860

; FILING DATE: 16-DEC-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Serafini, Andrew T.

; REGISTRATION NUMBER: 41,303

; REFERENCE/DOCKET NUMBER: 014643-00903005

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 576-0200

; TELEFAX: (415) 576-0300

; INFORMATION FOR SEQ ID NO: 272:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 12 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; SEQUENCE DESCRIPTION: SEQ ID NO: 272:

Query Match 100.0%; Score 7; DB 4; Length 12;  
Best Local Similarity 100.0%; Pred. No. 9.1e+03;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgtata 7  
|||||  
Db 4 TTTGATA 10

RESULT 5

US-08-166-664-5/c

; Sequence 5, Application US/08166664

; Patent No. 5646020

; GENERAL INFORMATION:

; APPLICANT: James A. McSwiggen

; APPLICANT: J. Anthony Mamone

; TITLE OF INVENTION: HAMMERHEAD RIBOZYMES FOR

; NUMBER OF SEQUENCES: 21

; CORRESPONDENCE ADDRESSES:

; ADDRESSEE: Lyon & Lyon

; STREET: 611 West Sixth Street

; CITY: Los Angeles

; STATE: California

; COUNTRY: USA

; ZIP: 90017

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: IBM P.C. DOS (Version 5.0)

; SOFTWARE: WordPerfect (Version 5.1)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/166,664

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/07/884,074

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Warburg, Richard J.

; REGISTRATION NUMBER: 32,327

; REFERENCE/DOCKET NUMBER: 197/062

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (213) 489-1600

; TELEFAX: (213) 955-0440

; TELEX: 67-3510

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 13

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; US-08-166-664-5

Query Match 100.0%; Score 7; DB 1; Length 13;

Best Local Similarity 100.0%; Pred. No. 9.1e+03;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgtata 7  
|||||

Db 10 TTTGATA 4

RESULT 6

US-08-271-880A-217/c

; Sequence 217, Application US/08271880A

; Patent No. 5693535

; GENERAL INFORMATION:

; APPLICANT: Kenneth G. Draper

; APPLICANT: Bharat Chovvita

; APPLICANT: James McSwiggen

; APPLICANT: Dan T. Stinchcomb

APPLICANT: James D. Thompson  
TITLE OF INVENTION: METHOD AND REAGENT FOR INHIBITING  
TITLE OF INVENTION: HUMAN IMMUNODEFICIENCY VIRUS  
TITLE OF INVENTION: REPLICATION  
NUMBER OF SEQUENCES: 232  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 MB  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: FASTSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/271,880A  
FILING DATE: July 7, 1994  
PRIOR APPLICATION DATA:  
PRIOR APPLICATION DATA: Including application  
PRIOR APPLICATION DATA: described below:  
APPLICATION NUMBER: 08/103,243  
FILING DATE: August 6, 1993  
APPLICATION NUMBER: 07/882,886  
FILING DATE: May 14, 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 206/116  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ. ID NO: 217:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-271-880A-217

Query Match 100.0%; Score 7; DB 1; Length 13;  
Best Local Similarity 100.0%; Pred. No. 9.1e+03;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ttgata 7  
1111111  
DB 10 TTGATA 4

RESULT 7  
US-08-910-408-217/C  
Sequence 217, Application US/08910408  
Patent No. 5972704  
GENERAL INFORMATION:  
APPLICANT: Kenneth G. Draper  
ADDRESSEE: Bharat Chowrira  
STREET: 633 West Fifth Street  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 MB  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: FASTSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/271,880A  
FILING DATE: July 7, 1994  
PRIOR APPLICATION DATA:  
PRIOR APPLICATION DATA: Including application  
PRIOR APPLICATION DATA: described below:  
APPLICATION NUMBER: 08/103,243  
FILING DATE: August 6, 1993  
APPLICATION NUMBER: 07/882,886  
FILING DATE: May 14, 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 206/116  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ. ID NO: 217:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-910-408-217

CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 MB  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: FASTSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/910,408  
FILING DATE:  
APPLICATION NUMBER: 08/271,880  
FILING DATE: July 7, 1994  
APPLICATION NUMBER: 08/103,243  
FILING DATE: August 6, 1993  
APPLICATION NUMBER: 07/882,886  
FILING DATE: May 14, 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 206/116  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ. ID NO: 217:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-910-408-217

Query Match 100.0%; Score 7; DB 2; Length 13;  
Best Local Similarity 100.0%; Pred. No. 9.1e+03;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ttgata 7  
1111111  
DB 10 TTGATA 4

RESULT 8  
US-09-249-215-217/C  
Sequence 217, Application US/09249215  
Patent No. 6159692  
GENERAL INFORMATION:  
APPLICANT: Kenneth G. Draper  
ADDRESSEE: Bharat Chowrira  
STREET: 633 West Fifth Street  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 MB  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0

SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/249,215  
FILING DATE: 12-Feb-1999  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/910,408  
FILING DATE: <Unknown>  
APPLICATION NUMBER: 08/103,243  
FILING DATE: August 6, 1993  
APPLICATION NUMBER: 07/882,886  
FILING DATE: May 14, 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 206/116  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 217:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 217:  
US-09-249-215-217

Query Match 100.0%; Score 7; DB 3; Length 13;  
Best Local Similarity 100.0%; Pred. No. 9.1e+03;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ttgata 7  
|||||  
Db 10 TTGATA 4

RESULT 9  
US-08-271-880A-212  
Sequence 212, Application US/08271880A  
Patent No. 5693535  
GENERAL INFORMATION:  
APPLICANT: Kenneth G. Draper  
APPLICANT: Bharat Chovitra  
APPLICANT: James McSwigen  
APPLICANT: Dan T. Stinchcomb  
APPLICANT: James D. Thompson  
TITLE OF INVENTION: METHOD AND REAGENT FOR INHIBITTING  
TITLE OF INVENTION: HUMAN IMMUNODEFICIENCY VIRUS  
NUMBER OF SEQUENCES: 232  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
STREET: Suite 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/271,880A  
FILING DATE: July 7, 1994  
PRIOR APPLICATION DATA:  
PRIOR APPLICATION DATA: including application  
PRIOR APPLICATION DATA: described below:  
APPLICATION NUMBER: 08/103,243 two

FILING DATE: August 6, 1993  
APPLICATION NUMBER: 07/882,886  
FILING DATE: May 14, 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 206/116  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 212:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-271-880A-212

Query Match 100.0%; Score 7; DB 1; Length 14;  
Best Local Similarity 100.0%; Pred. No. 9.1e+03;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ttgata 7  
|||||  
Db 3 TTGATA 9

RESULT 10  
US-08-910-408-212  
Sequence 212, Application US/08910408  
Patent No. 5972704  
GENERAL INFORMATION:  
APPLICANT: Kenneth G. Draper  
APPLICANT: Bharat Chovitra  
APPLICANT: James McSwigen  
APPLICANT: Dan T. Stinchcomb  
APPLICANT: James D. Thompson  
TITLE OF INVENTION: METHOD AND REAGENT FOR INHIBITTING  
TITLE OF INVENTION: HUMAN IMMUNODEFICIENCY VIRUS  
NUMBER OF SEQUENCES: 232  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
STREET: Suite 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/910,408  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/271,880  
FILING DATE: July 7, 1994  
APPLICATION NUMBER: 08/103,243  
FILING DATE: August 6, 1993  
APPLICATION NUMBER: 07/882,886  
FILING DATE: May 14, 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 206/116  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600

TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 212:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-910-408-212

Query Match 100.0%; Score 7; DB 2; Length 14;  
Best Local Similarity 100.0%; Pred. No. 9.1e+03;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ttgata 7  
|||||  
DB 3 TTGGATA 9

RESULT 11  
US-09-249-215-212  
Sequence 212, Application US/09249215  
Patent No. 6159692  
GENERAL INFORMATION:  
APPLICANT: Kenneth G. Draper  
Bharat Chowrita  
James MCSwlggen  
Dan T. Stinchcomb  
James D. Thompson

TITLE OF INVENTION: METHOD AND REAGENT FOR INHIBITING  
HUMAN IMMUNODEFICIENCY VIRUS  
REPLICATION

NUMBER OF SEQUENCES: 232  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
Suite 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
storage

COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: FASTSEQ Version 1.5

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/249,215  
FILING DATE: 12-Feb-1999  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/910,408

FILING DATE: <unknown>  
APPLICATION NUMBER: 08/103,243  
FILING DATE: August 6, 1993  
APPLICATION NUMBER: 07/882,886  
FILING DATE: May 14, 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 206/116

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 212:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 212:

US-09-249-215-212

Query Match 100.0%; Score 7; DB 3; Length 14;  
Best Local Similarity 100.0%; Pred. No. 9.1e+03;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ttgata 7  
|||||  
DB 3 TTGGATA 9

RESULT 12  
US-08-373-124A-90/c  
Sequence 90, Application US/08373124A  
Patent No. 5646042

GENERAL INFORMATION:  
APPLICANT: Stinchcomb, Dan T.  
APPLICANT: Draper, Kenneth  
APPLICANT: MCSwlggen, James  
APPLICANT: Jarvis, Thale  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR  
TREATMENT OF RESPONOSIS AND  
TITLE OF INVENTION: CANCER USING RIBOZYMES  
NUMBER OF SEQUENCES: 2627  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
Suite 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/373,124A  
FILING DATE: January 13, 1995

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/245,466  
FILING DATE: May 18, 1994  
APPLICATION NUMBER: 08/192,943  
FILING DATE: February 7, 1994  
APPLICATION NUMBER: 07/987,132  
FILING DATE: December 7, 1992  
APPLICATION NUMBER: 07/936,422  
FILING DATE: August 26, 1992

ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 209/035  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 90:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-373-124A-90

Query Match 100.0%; Score 7; DB 1; Length 15;  
Best Local Similarity 100.0%; Pred. No. 9.1e+03;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ttgata 7

Db 11 TTTGATA 5

## RESULT 13

US-08-291-932A-129  
Sequence 129, Application US/08291932A  
Patent No. 5658780

## GENERAL INFORMATION:

APPLICANT: Stinchcomb, Dan T.  
APPLICANT: Draper, Kenneth G.  
APPLICANT: McSwiggen, James  
TITLE OF INVENTION: RIBOZYME TREATMENT OF  
DISEASES OR CONDITIONS  
TITLE OF INVENTION: RELATED TO LEVELS OF  
TITLE OF INVENTION: NF-KB  
NUMBER OF SEQUENCES: 830  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071-2066

## COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/291,932A  
FILING DATE: August 15, 1994

## CLASSIFICATION: 514

PRIOR APPLICATION DATA: Including Application  
PRIOR APPLICATION DATA: described below:  
APPLICATION NUMBER: 08/245,466  
FILING DATE: May 18, 1994

Two

APPLICATION NUMBER: 07/987,132  
FILING DATE: December 7, 1992

## ATTORNEY/AGENT INFORMATION:

NAME: Wardburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 208/157  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440

## TELEX: 67-3510

## INFORMATION FOR SEQ ID NO: 129:

## SEQUENCE CHARACTERISTICS:

LENGTH: 15 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-291-932A-129

## Query Match

Best Local Similarity 100.0%; Score 7; DB 1; Length 15;  
Matches 3; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ttgtgata 7

Db 9 UUGCAUA 15

## RESULT 14

US-08-291-932A-130  
Sequence 130, Application US/08291932A  
Patent No. 5658780  
GENERAL INFORMATION:

APPLICANT: Stinchcomb, Dan T.  
APPLICANT: Draper, Kenneth G.  
APPLICANT: McSwiggen, James

TITLE OF INVENTION: RIBOZYME TREATMENT OF  
DISEASES OR CONDITIONS  
TITLE OF INVENTION: RELATED TO LEVELS OF  
TITLE OF INVENTION: NF-KB  
NUMBER OF SEQUENCES: 830  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071-2066

## COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/291,932A  
FILING DATE: August 15, 1994

## CLASSIFICATION: 514

PRIOR APPLICATION DATA: Including application  
PRIOR APPLICATION DATA: described below:  
APPLICATION NUMBER: 08/245,466  
FILING DATE: May 18, 1994

Two

APPLICATION NUMBER: 07/987,132  
FILING DATE: December 7, 1992

## ATTORNEY/AGENT INFORMATION:

NAME: Wardburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 208/157  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440

## TELEX: 67-3510

## INFORMATION FOR SEQ ID NO: 130:

## SEQUENCE CHARACTERISTICS:

LENGTH: 15 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-291-932A-130

## Query Match

Best Local Similarity 100.0%; Score 7; DB 1; Length 15;  
Matches 3; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ttgtgata 7

Db 7 UUGCAUA 13

## RESULT 15

US-08-291-932A-131  
Sequence 131, Application US/08291932A  
Patent No. 5658780

## GENERAL INFORMATION:

APPLICANT: Stinchcomb, Dan T.  
APPLICANT: Draper, Kenneth G.  
APPLICANT: McSwiggen, James  
TITLE OF INVENTION: RIBOZYME TREATMENT OF  
DISEASES OR CONDITIONS  
TITLE OF INVENTION: RELATED TO LEVELS OF  
TITLE OF INVENTION: NF-KB  
NUMBER OF SEQUENCES: 830  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon  
 STREET: 633 West Fifth Street  
 STREET: Suite 4700  
 CITY: Los Angeles  
 STATE: California  
 COUNTRY: U.S.A.  
 ZIP: 90071-2066  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
 MEDIUM TYPE: storage  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: IBM P.C. DOS 5.0  
 SOFTWARE: Word Perfect 5.1  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/291,932A  
 FILING DATE: August 15, 1994  
 CLASSIFICATION: 514  
 PRIOR APPLICATION DATA: including application  
 PRIOR APPLICATION DATA: described below:  
 APPLICATION NUMBER: 08/245,466  
 FILING DATE: May 18, 1994  
 APPLICATION NUMBER: 07/987,132  
 FILING DATE: December 7, 1992  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Wardburg, Richard J.  
 REGISTRATION NUMBER: 32,327  
 REFERENCE/DOCKET NUMBER: 208/157  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (213) 489-1600  
 TELEFAX: (213) 955-0440  
 TELEEX: 67-3510  
 INFORMATION FOR SEQ ID NO: 131:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 15 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 US-08-291-932A-131

Two

Query Match 100.0%; Score 7; DB 1; Length 15;  
 Best Local Similarity 42.9%; Pred. NO. 9.1e+03;  
 Matches 3; Conservative 4; Mismatches 0; Indels 0;  
 Gaps 0;  
 OY 1 ttgata 7  
 Db 6 UUGAUA 12

Search completed: December 15, 2001, 03:22:47  
 Job time: 7562 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 02:33:41 : Search time 4441.54 Seconds  
(without alignments)  
16.936 Million cell updates/sec

Title: US-09-380-826a-5

Perfect score: 7  
Sequence: 1 ttgata 7

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

EST: \*  
1: em\_estfun: \*  
2: em\_esthum: \*  
3: em\_estin: \*  
4: em\_estom: \*  
5: em\_estpl: \*  
6: em\_estba: \*  
7: em\_estro: \*  
8: em\_estov: \*  
9: em\_hlc: \*  
10: gb\_estf: \*  
11: gb\_estc: \*  
12: gb\_hlc: \*  
13: gb\_gss: \*  
14: em\_gss\_fun: \*  
15: em\_gss\_hum: \*  
16: em\_gss\_inv: \*  
17: em\_gss\_pln: \*  
18: em\_gss\_pro: \*  
19: em\_gss\_rtd: \*  
20: em\_gss\_vrt: \*  
21: em\_gss\_other: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	7	100.0	19	AZ641145	AZ641145 1M0503F04
C 2	7	100.0	20	AZ323820	AZ323820 1M0045A01
C 3	7	100.0	20	AZ819886	AZ819886 2M0091H16
C 4	7	100.0	20	AZ977536	AZ977536 2M0253B04
C 5	7	100.0	22	AZ329387	AZ329387 1M0053C07
C 6	7	100.0	22	AZ502952	AZ502952 1M0342C04
C 7	7	100.0	22	AZ828775	AZ828775 2M0106B06
C 8	7	100.0	23	AZ345476	AZ345476 1M0080P11
C 9	7	100.0	23	AZ789315	AZ789315 2M0036M22
C 10	7	100.0	23	AZ840208	AZ840208 2M0136H12
C 11	7	100.0	23	AZ862305	AZ862305 2M0169N22
C 12	7	100.0	23	TA144D040	TA144D040 T. brucei

C	13	7	100.0	24	13	AZ379215	AZ379215 1M0134M14
C 14	7	100.0	24	13 <td>AZ480650<td>AZ480650 1M0302P17</td></td>	AZ480650 <td>AZ480650 1M0302P17</td>	AZ480650 1M0302P17	
C 15	7	100.0	24	13 <td>AZ663958<td>AZ663958 1M0543R19</td></td>	AZ663958 <td>AZ663958 1M0543R19</td>	AZ663958 1M0543R19	
C 16	7	100.0	24	13 <td>AZ794696<td>AZ794696 2M0048C04</td></td>	AZ794696 <td>AZ794696 2M0048C04</td>	AZ794696 2M0048C04	
C 17	7	100.0	25	10 <td>AU008929<td>AU008929 AU008929</td></td>	AU008929 <td>AU008929 AU008929</td>	AU008929 AU008929	
C 18	7	100.0	25	13 <td>AZ404619<td>AZ404619 1M0173L20</td></td>	AZ404619 <td>AZ404619 1M0173L20</td>	AZ404619 1M0173L20	
C 19	7	100.0	25	13 <td>AZ592455<td>AZ592455 1M0403H19</td></td>	AZ592455 <td>AZ592455 1M0403H19</td>	AZ592455 1M0403H19	
C 20	7	100.0	26	10 <td>AM333097<td>AM333097 S17C3 AGS</td></td>	AM333097 <td>AM333097 S17C3 AGS</td>	AM333097 S17C3 AGS	
C 21	7	100.0	27	11 <td>D18735<td>D18735 MUSGS01797</td></td>	D18735 <td>D18735 MUSGS01797</td>	D18735 MUSGS01797	
C 22	7	100.0	27	13 <td>AZ328549<td>AZ328549 1M0052D24</td></td>	AZ328549 <td>AZ328549 1M0052D24</td>	AZ328549 1M0052D24	
C 23	7	100.0	27	13 <td>AZ335603<td>AZ335603 1M0065E13</td></td>	AZ335603 <td>AZ335603 1M0065E13</td>	AZ335603 1M0065E13	
C 24	7	100.0	27	13 <td>AZ836137<td>AZ836137 2M0130M13</td></td>	AZ836137 <td>AZ836137 2M0130M13</td>	AZ836137 2M0130M13	
C 25	7	100.0	27	13 <td>AZ966573<td>AZ966573 2M0237B12</td></td>	AZ966573 <td>AZ966573 2M0237B12</td>	AZ966573 2M0237B12	
C 26	7	100.0	28	13 <td>AZ850027<td>AZ850027 2M0151P09</td></td>	AZ850027 <td>AZ850027 2M0151P09</td>	AZ850027 2M0151P09	
C 27	7	100.0	29	13 <td>AZ966795<td>AZ966795 2M0237P23</td></td>	AZ966795 <td>AZ966795 2M0237P23</td>	AZ966795 2M0237P23	
C 28	7	100.0	29	13 <td>TA97B05P<td>TA97B05P T. brucei</td></td>	TA97B05P <td>TA97B05P T. brucei</td>	TA97B05P T. brucei	
C 29	7	100.0	30	10 <td>AU007442<td>AU007442 AU007442</td></td>	AU007442 <td>AU007442 AU007442</td>	AU007442 AU007442	
C 30	7	100.0	30	11 <td>C20899<td>C20899 HUMGS000497</td></td>	C20899 <td>C20899 HUMGS000497</td>	C20899 HUMGS000497	
C 31	7	100.0	30	11 <td>BE911076<td>BE911076 601662175</td></td>	BE911076 <td>BE911076 601662175</td>	BE911076 601662175	
C 32	7	100.0	30	13 <td>AZ601864<td>AZ601864 1M0420P16</td></td>	AZ601864 <td>AZ601864 1M0420P16</td>	AZ601864 1M0420P16	
C 33	7	100.0	30	13 <td>TA118C09P<td>TA118C09P T. brucei</td></td>	TA118C09P <td>TA118C09P T. brucei</td>	TA118C09P T. brucei	
C 34	7	100.0	31	10 <td>A1140482<td>A1140482 Cw82d12.S</td></td>	A1140482 <td>A1140482 Cw82d12.S</td>	A1140482 Cw82d12.S	
C 35	7	100.0	31	10 <td>A1973656<td>A1973656 sc07h11.Y</td></td>	A1973656 <td>A1973656 sc07h11.Y</td>	A1973656 sc07h11.Y	
C 36	7	100.0	31	13 <td>AZ591446<td>AZ591446 1M0401J05</td></td>	AZ591446 <td>AZ591446 1M0401J05</td>	AZ591446 1M0401J05	
C 37	7	100.0	31	13 <td>AZ799349<td>AZ799349 2M0056G12</td></td>	AZ799349 <td>AZ799349 2M0056G12</td>	AZ799349 2M0056G12	
C 38	7	100.0	32	10 <td>AU009849<td>AU009849 AU009849</td></td>	AU009849 <td>AU009849 AU009849</td>	AU009849 AU009849	
C 39	7	100.0	32	11 <td>D18230<td>D18230 MUSGS00509</td></td>	D18230 <td>D18230 MUSGS00509</td>	D18230 MUSGS00509	
C 40	7	100.0	32	13 <td>AZ447488<td>AZ447488 1M0244K07</td></td>	AZ447488 <td>AZ447488 1M0244K07</td>	AZ447488 1M0244K07	
C 41	7	100.0	33	11 <td>R98634<td>R98634 Yq63h06.r1</td></td>	R98634 <td>R98634 Yq63h06.r1</td>	R98634 Yq63h06.r1	
C 42	7	100.0	33	13 <td>AZ308864<td>AZ308864 1M0012M13</td></td>	AZ308864 <td>AZ308864 1M0012M13</td>	AZ308864 1M0012M13	
C 43	7	100.0	33	13 <td>AZ411435<td>AZ411435 1M0184C05</td></td>	AZ411435 <td>AZ411435 1M0184C05</td>	AZ411435 1M0184C05	
C 44	7	100.0	33	13 <td>AZ768977<td>AZ768977 1M0569B14</td></td>	AZ768977 <td>AZ768977 1M0569B14</td>	AZ768977 1M0569B14	
C 45	7	100.0	34	13 <td>TA164G090<td>TA164G090 T. brucei</td></td>	TA164G090 <td>TA164G090 T. brucei</td>	TA164G090 T. brucei	

#### ALIGNMENTS

RESULT 1  
LOCUS AZ641145/c 19 bp DNA  
DEFINITION IM0503F04R Mouse 10kb plasmid UNGC1M library Mus musculus genomic  
clone UNGC1M0503F04 R, DNA sequence.  
ACCESSION AZ641145  
VERSION AZ641145.1 GI:11764838  
KEYWORDS GSS.  
SOURCE house mouse.  
ORGANISM Mus musculus  
REFERENCE 1 (bases 1 to 19)  
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamli,C.,  
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,  
'M., Rose,M., Rose,R., Stokes,R., Tingley,A., von Niederhausern,A.  
and Wright,D., Weiss,R.  
TITLE Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts  
JOURNAL Unpublished (2000)  
COMMENT Contact: Robert B. Weiss  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0503 row: F column: 04  
Seq primer: CACACAGCAACACGCTATGACC  
Class: plasmid ends  
High quality sequence stop: 19.  
Location/Qualifiers  
1. 19

FEATURES  
source

/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUGC1M0503F04"  
/clone.lib="Mouse 10kb plasmid UUGC1M library"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
/note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g1473211419b1AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT  
ORIGIN  
7 a 5 c 2 g 5 t

Query Match 100.0%; Score 7; DB 13; Length 19;  
Best local Similarity 100.0%; Pred. No. 4.2e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ttgtgata 7  
|||||||  
Db 19 TTTGATTA 13

RESULT 2  
A2323820 20 bp DNA GSS 29-SEP-2000  
LOCUS 1M0045A01R Mouse 10kb plasmid UUGC1M library Mus musculus genomic  
DEFINITION clone UUGC1M0045A01 R, DNA sequence.  
ACCESSION A2323820  
VERSION A2323820.1 GI:10378917  
KEYWORDS GSS.  
SOURCE house mouse.  
ORGANISM Mus musculus.  
REFERENCE 1 (bases 1 to 20)  
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A., and Wright,D., Weiss,R.  
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts  
JOURNAL Unpublished (2000)  
COMMENT Contact: Robert B. Weiss  
University of Utah Genome Center  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0045 row: A column: 01  
Seq primer: CACACAGAAACAGCATATGACC  
Class: plasmid ends  
High quality sequence stop: 20.  
Location/Qualifiers

source  
1. 20  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUGC1M0045A01"  
/clone.lib="Mouse 10kb plasmid UUGC1M library"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
/note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g1473211419b1AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT  
ORIGIN  
10 a 4 c 3 g 3 t

Query Match 100.0%; Score 7; DB 13; Length 20;  
Best local Similarity 100.0%; Pred. No. 4.2e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ttgtgata 7  
|||||||  
Db 11 TTTGATA 5

RESULT 3  
A2819886 20 bp DNA GSS 20-FEB-2001  
LOCUS 2M0091H16R Mouse 10kb plasmid UUGC1M library Mus musculus genomic  
DEFINITION clone UUGC2M0091H16 R, DNA sequence.  
ACCESSION A2819886  
VERSION A2819886.1 GI:12989794  
KEYWORDS GSS.  
SOURCE house mouse.  
ORGANISM Mus musculus.  
REFERENCE 1 (bases 1 to 20)  
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A., and Wright,D., Weiss,R.  
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts  
JOURNAL Unpublished (2000)  
COMMENT Contact: Robert B. Weiss  
University of Utah Genome Center  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0091 row: H column: 16  
Seq primer: CACACAGAAACAGCATATGACC  
Class: plasmid ends  
High quality sequence stop: 20.

FEATURES  
source

Location/Qualifiers  
1. .20  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="U06C2M053B04"  
/clone\_1lb="Mouse 10kb plasmid U06C2M library"  
/sex="Male"  
/lab\_host="E. coli strain XL10-Gold, T1-resistant, F-"  
/note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g114732114|gblAF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT  
ORIGIN  
7 a 3 c 4 g 6 t

Query Match  
Best Local Similarity 100.0%; Score 7; DB 13; Length 20;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ttgata 7  
|||||  
Db 10 TTTGATA 16

RESULT 4  
A2977536 20 bp DNA GSS 27-APR-2001  
LOCUS 2M053B04R Mouse 10kb plasmid U06C2M library Mus musculus genomic  
DEFINITION clone U06C2M053B04 R, DNA sequence.  
ACCESSION A2977536  
VERSION A2977536.1 GI:13848763  
KEYWORDS GSS.  
SOURCE house mouse.  
ORGANISM Mus musculus.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 20)  
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Roge,R., Stokes,R., Tinney,A., von Niederhausern,A. and Wright,D., Weiss,R.  
Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts  
Unpublished (2000)  
Contact: Robert B. Weiss  
University of Utah Genome Center  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT 84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: [ddunn@genetics.utah.edu](mailto:ddunn@genetics.utah.edu)  
Insert length: 10000 Std Error: 0.00  
Plate: 0253 row: B column: 04  
Seq primer: CACACAGGAACACGTATGACC  
Class: plasmid ends

JOURNAL  
COMMENT

FEATURES  
source

High quality sequence stop: 20.  
Location/Qualifiers  
1. .20  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="U06C2M053B04"  
/clone\_1lb="Mouse 10kb plasmid U06C2M library"  
/sex="Female"  
/lab\_host="E. coli strain XL10-Gold, T1-resistant, F-"  
/note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (female) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g114732114|gblAF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT  
ORIGIN  
4 a 5 c 2 g 9 t

Query Match  
Best Local Similarity 100.0%; Score 7; DB 13; Length 20;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ttgata 7  
|||||  
Db 7 TTTGATA 13

RESULT 5  
A2329387 22 bp DNA GSS 29-SEP-2000  
LOCUS 1M0053C07R Mouse 10kb plasmid U06C1M library Mus musculus genomic  
DEFINITION clone U06C1M0053C07 R, DNA sequence.  
ACCESSION A2329387  
VERSION A2329387.1 GI:10390050  
KEYWORDS GSS.  
SOURCE house mouse.  
ORGANISM Mus musculus.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 22)  
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Roge,R., Stokes,R., Tinney,A., von Niederhausern,A. and Wright,D., Weiss,R.  
Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts  
Unpublished (2000)  
Contact: Robert B. Weiss  
University of Utah Genome Center  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT 84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: [ddunn@genetics.utah.edu](mailto:ddunn@genetics.utah.edu)  
Insert length: 10000 Std Error: 0.00  
Plate: 0053 row: C column: 07  
Seq primer: CACACAGGAACACGTATGACC

JOURNAL  
COMMENT

Class: plasmid ends  
High quality sequence stop: 22.  
Location/Qualifiers  
source

1. 22  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUGC1M0053C07"  
/clone\_lib="Mouse 10kb plasmid UUGC1M library"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
/note="Vector: PMD42ny; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g11473211419b1AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 8 a 4 c 3 g 7 t  
ORIGIN

Query Match 100.0%; Score 7; DB 13; Length 22;  
Best Local Similarity 100.0%; Pred. No. 4.2e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ttgtgata 7  
|||||  
Db 21 TTGTGATA 15

RESULT 6  
AZ502952/c 22 bp DNA GSS 05-OCT-2000  
LOCUS 1M0342C04R Mouse 10kb plasmid UUGC1M library Mus musculus genomic  
DEFINITION clone UUGC1M0342C04 R, DNA sequence.  
ACCESSION AZ502952  
VERSION AZ502952.1 GI:10684268  
KEYWORDS GSS.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 22)  
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,  
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly  
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.  
and Wright,D.,Weiss,R.  
Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts  
Unpublished (2000)  
Contact: Robert B. Weiss  
University of Utah Genome Center  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0342 row: C column: 04

Seq primer: CACACAGAAACAGCTATGACC  
Class: plasmid ends  
High quality sequence stop: 22.  
Location/Qualifiers  
source

1. 22  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUGC1M0342C04"  
/clone\_lib="Mouse 10kb plasmid UUGC1M library"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
/note="Vector: PMD42ny; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g11473211419b1AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 10 a 5 c 1 g 6 t  
ORIGIN

Query Match 100.0%; Score 7; DB 13; Length 22;  
Best Local Similarity 100.0%; Pred. No. 4.2e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ttgtgata 7  
|||||  
Db 9 TTGTGATA 3

RESULT 7  
AZ828775/c 22 bp DNA GSS 20-FEB-2001  
LOCUS 2M0106B06F Mouse 10kb plasmid UUGC1M library Mus musculus genomic  
DEFINITION clone UUGC2M0106B06 F, DNA sequence.  
ACCESSION AZ828775  
VERSION AZ828775.1 GI:12998683  
KEYWORDS GSS.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 22)  
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,  
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly  
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.  
and Wright,D.,Weiss,R.  
Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts  
Unpublished (2000)  
Contact: Robert B. Weiss  
University of Utah Genome Center  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00

Plate: 0106 row: B column: 06  
Seq primer: CGTGTAAACGACGCCACGT  
Class: plasmid ends  
High quality sequence stop: 22.  
Location/Qualifiers

FEATURES  
source

```
1. 22
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="U06C2M0106B06"
/clone_lib="Mouse 10kb plasmid U06C1M library"
/sex="male"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PMD42 (g114732114[gblAF129072.1], a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
```

BASE COUNT  
ORIGIN

9 a 7 c 2 g 4 t

Query Match 100.0%; Score 7; DB 13; Length 22;  
Best Local Similarity 100.0%; Pred. No. 4.2e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ttgtata 7  
|||||

Db 21 TTTGATA 15

RESULT 8  
A2345476/c 23 bp DNA GSS 29-SEP-2000  
LOCUS 1M0080P11F Mouse 10kb plasmid U06C1M library Mus musculus genomic  
DEFINITION clone U06C1M0080P11 F, DNA sequence.

ACCESSION A2345476  
VERSION A2345476.1 GI:10424713  
KEYWORDS GSS.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 23)  
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,  
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly  
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.  
and Wright,D., Weiss,R.  
Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts  
Unpublished (2000)  
Contact: Robert B. Weiss  
University of Utah Genome Center  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel.: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu

JOURNAL COMMENT

Insert Length: 10000 Std Error: 0.00  
Plate: 0080 row: P column: 11  
Seq primer: CGTGTAAACGACGCCACGT  
Class: plasmid ends  
High quality sequence stop: 23.  
Location/Qualifiers

FEATURES  
source

```
1. 23
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="U06C1M0080P11"
/clone_lib="Mouse 10kb plasmid U06C1M library"
/sex="male"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PMD42 (g114732114[gblAF129072.1], a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
```

BASE COUNT  
ORIGIN

12 a 1 c 0 g 10 t

Query Match 100.0%; Score 7; DB 13; Length 23;  
Best Local Similarity 100.0%; Pred. No. 4.2e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ttgtata 7  
|||||

Db 23 TTTGATA 17

RESULT 9  
A2789315/c 23 bp DNA GSS 16-FEB-2001  
LOCUS 2M0036W22R Mouse 10kb plasmid U06C1M library Mus musculus genomic  
DEFINITION clone U06C2M0036W22 R, DNA sequence.

ACCESSION A2789315  
VERSION A2789315.1 GI:12929986  
KEYWORDS GSS.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 23)  
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,  
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly  
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.  
and Wright,D., Weiss,R.  
Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts  
Unpublished (2000)  
Contact: Robert B. Weiss  
University of Utah Genome Center  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel.: 801 585 5606  
Fax: 801 585 7177

JOURNAL COMMENT

Email: ddunn@genetics.utah.edu  
 Insert Length: 10000 Std Error: 0.00  
 Plate: 0036 row: M column: 22  
 Seq primer: CACACAGAAACAGCTATGACC  
 Class: plasmid ends  
 High quality sequence stop: 23.  
 Location/Qualifiers

# FEATURES

## SOURCE

1. 23  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="UUCG2M0136H12"  
 /clone\_11b="Mouse 10kb plasmid UUCG1M library"  
 /sex="Male"  
 /lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
 /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g11473211419b/AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

## BASE COUNT

9 a 4 c 4 g 6 t

## ORIGIN

Query Match 100.0%; Score 7; DB 13; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 4.2e+05;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ttgtgata 7  
 |||||  
 Db 22 TTTGATA 16

## RESULT 10

AZ840208 23 bp DNA GSS 20-FEB-2001  
 LOCUS 2M0136H12R Mouse 10kb plasmid UUCG1M library Mus musculus genomic  
 DEFINITION clone UUCG2M0136H12 R, DNA sequence.

ACCESSION AZ840208.1 GI:13010116

VERSION GSS.  
 KEYWORDS house mouse.  
 SOURCE Mus musculus

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 23)

Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,  
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,  
 M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.  
 and Wright,D.,Weiss,R.

Mouse whole genome scaffolding with paired end reads from 10kb

TITLE Plasmid Inserts  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Robert B. Weiss  
 University of Utah Genome Center  
 University of Utah  
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
 84112, USA  
 Tel: 801 585 5606

Fax: 801 585 7177  
 Email: ddunn@genetics.utah.edu  
 Insert Length: 10000 Std Error: 0.00  
 Plate: 0136 row: H column: 12  
 Seq primer: CACACAGAAACAGCTATGACC  
 Class: plasmid ends  
 High quality sequence stop: 23.  
 Location/Qualifiers

# FEATURES

## SOURCE

1. 23  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="UUCG2M0136H12"  
 /clone\_11b="Mouse 10kb plasmid UUCG1M library"  
 /sex="Male"  
 /lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
 /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g11473211419b/AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

## BASE COUNT

8 a 7 c 3 g 5 t

## ORIGIN

Query Match 100.0%; Score 7; DB 13; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 4.2e+05;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ttgtgata 7  
 |||||  
 Db 17 TTTGATA 11

## RESULT 11

AZ862305 23 bp DNA GSS 21-FEB-2001  
 LOCUS 2M0169N22F Mouse 10kb plasmid UUCG1M library Mus musculus genomic  
 DEFINITION clone UUCG2M0169N22 F, DNA sequence.

ACCESSION AZ862305

VERSION AZ862305.1 GI:13059475

## KEYWORDS

GSS.  
 house mouse.  
 SOURCE Mus musculus

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 23)

Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,  
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,  
 M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.  
 and Wright,D.,Weiss,R.

Mouse whole genome scaffolding with paired end reads from 10kb

TITLE Plasmid Inserts  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Robert B. Weiss  
 University of Utah Genome Center  
 University of Utah  
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
 84112, USA

Tel: 801 585 5606  
 Fax: 801 585 7177  
 Email: ddunn@genetics.utah.edu  
 Insert Length: 10000 Std Error: 0.00  
 Plate: 0169 row: N column: 22  
 Seq primer: CGTTGTAACGACGCCACG  
 Class: plasmid ends  
 High quality sequence stop: 23.  
 Location/Qualifiers

## FEATURES

source

1. 23  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="UUCG2M0169N22"  
 /clone\_1lb="Mouse 10kb plasmid UUCG1M library"  
 /sex="Male"  
 /lab\_host="E. Coli strain XL10-Gold, TI-resistant, F-"  
 /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g11473211419b1AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

## BASE COUNT

7 a 7 c 1 g 8 t

## ORIGIN

Query Match 100.0%; Score 7; DB 13; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 4.2e+05;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 ttgtgata 7  
 |||||  
 Db 17 TTTGATA 11

## RESULT 12

TA144D040

LOCUS TA144D040 23 bp DNA GSS 13-DEC-2000  
 DEFINITION T. brucei sheared genomic DNA clone 144d04, reverse sequence,  
 genomic survey sequence.

ACCESSION AL466870  
 VERSION AL466870.1 GI:11836225  
 GSS.

KEYWORDS Trypanosoma brucei.  
 SOURCE Trypanosoma brucei.  
 ORGANISM Trypanosoma brucei.  
 Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae;  
 Trypanosoma.

## REFERENCE

AUTHORS

1 (bases 1 to 23)  
 Hall, N., Bowman, S., Lennard, N.J., Doggett, J., Atkin, R.,  
 Chillingworth, C., Ormond, D., Harris, B., El-Sayed, N., Hou, L.,  
 Melville, S.E., Rajandream, M.A. and Barrell, B.G.

## TITLE

JOURNAL

Direct Submission  
 Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing  
 project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,  
 Cambridge CB10 1SA, Email: barrell@sanger.ac.uk and  
 nhle@sanger.ac.uk

## COMMENT

Constructed at the Institute for Genomic Research (TIGR),  
 Rockville, MD. Genomic DNA isolated from a cloned population of  
 Trypanosoma brucei (TREU927/4 G9rat 10.1) was mechanically sheared

to give a tight size distribution (4 kb). The v + i method used for the library construction is described in detail in Smith, H. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999).  
 Email: nelsayed@tigr.org  
 Details of T. brucei sequencing at the Sanger Centre are available at <http://www.sanger.ac.uk/projects/T-brucei/>.

## FEATURES

source

1. 23  
 /organism="Trypanosoma brucei"  
 /strain="TREU927"  
 /db\_xref="taxon:5691"  
 /clone="144d04"  
 /clone\_1lb="Mouse 10kb plasmid UUCG1M library"  
 /sex="Male"  
 /lab\_host="E. Coli strain XL10-Gold, TI-resistant, F-"  
 /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource

## BASE COUNT

3 a 6 c 5 g 9 t

## ORIGIN

Query Match 100.0%; Score 7; DB 13; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 4.2e+05;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 ttgtgata 7  
 |||||  
 Db 11 TTTGATA 17

## RESULT 13

AZ379215/c

LOCUS AZ379215 24 bp DNA GSS 02-OCT-2000  
 DEFINITION IM0134M14F Mouse 10kb plasmid UUCG1M library Mus musculus genomic  
 clone UUCG1M0134M14 F, DNA sequence.

ACCESSION AZ379215  
 VERSION AZ379215.1 GI:10492915  
 KEYWORDS GSS.  
 SOURCE house mouse.  
 ORGANISM Mus musculus.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

## REFERENCE

AUTHORS

1 (bases 1 to 24)  
 Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C.,  
 Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,  
 M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.  
 and Wright, D., Weiss, R.

## TITLE

JOURNAL

Mouse whole genome scaffolding with paired end reads from 10kb  
 plasmid inserts  
 Unpublished (2000)  
 Contact: Robert B. Weiss  
 University of Utah  
 University of Utah  
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
 84112, USA

Tel: 801 585 5606  
 Fax: 801 585 7177  
 Email: ddunn@genetics.utah.edu  
 Insert Length: 10000 Std Error: 0.00  
 Plate: 0134 row: M column: 14  
 Seq primer: CGTTGTAACGACGCCACG  
 Class: plasmid ends  
 High quality sequence stop: 24.  
 Location/Qualifiers

## FEATURES

source

1. 24  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="UUCG1M0134M14"  
 /clone\_1lb="Mouse 10kb plasmid UUCG1M library"  
 /sex="Male"  
 /lab\_host="E. Coli strain XL10-Gold, TI-resistant, F-"  
 /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g11473211419b|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 8 a 4 c 4 g 8 t  
ORIGIN

Query Match 100.0%; Score 7; DB 13; Length 24;  
Best Local Similarity 100.0%; Pred. No. 4.2e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgata 7  
1111111  
DB 12 TTGTGATA 6

RESULT 14  
AZ480650 24 bp DNA GSS 04-OCT-2000  
LOCUS 1M0302P17F Mouse 10kb plasmid UUGC1M library Mus musculus genomic  
DEFINITION clone UUGC1M0302P17 F, DNA sequence.  
ACCESSION AZ480650  
VERSION AZ480650.1 GI:10641715  
KEYWORDS GSS.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 24)  
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,  
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly  
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.  
and Wright,D., Weiss,R.  
Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts

JOURNAL Unpublished (2000)  
COMMENT Contact: Robert B. Weiss  
University of Utah Genome Center  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0302 row: P column: 17  
Seq primer: CGTGTAAACAGCGCCAGT  
Class: plasmid ends  
High quality sequence stop: 24.

FEATURES  
source 1..24  
location/Qualifiers  
1..24  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUGC1M0302P17"  
/clone\_11b="Mouse 10kb plasmid UUGC1M library"  
/sex="Male"  
/lab\_host="E. coli strain XL10-Gold, T1-resistant, F-"  
/note="Vector: PMD42nv; Purified genomic DNA from M.  
musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g11473211419b|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 11 a 4 c 1 g 8 t  
ORIGIN

Query Match 100.0%; Score 7; DB 13; Length 24;  
Best Local Similarity 100.0%; Pred. No. 4.2e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgata 7  
1111111  
DB 8 TTGTGATA 2

RESULT 15  
AZ663958 24 bp DNA GSS 14-DEC-2000  
LOCUS 1M0543E19R Mouse 10kb plasmid UUGC1M library Mus musculus genomic  
DEFINITION clone UUGC1M0543E19 R, DNA sequence.  
ACCESSION AZ663958  
VERSION AZ663958.1 GI:11801104  
KEYWORDS GSS.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 24)  
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,  
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly  
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.  
and Wright,D., Weiss,R.  
Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts

JOURNAL Unpublished (2000)  
COMMENT Contact: Robert B. Weiss  
University of Utah Genome Center  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0543 row: E column: 19  
Seq primer: CACACAGAAACAGCTATGACC  
Class: plasmid ends  
High quality sequence stop: 24.

FEATURES  
source 1..24  
location/Qualifiers  
1..24  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUGC1M0543E19"  
/clone\_11b="Mouse 10kb plasmid UUGC1M library"  
/sex="Male"  
/lab\_host="E. coli strain XL10-Gold, T1-resistant, F-"  
/note="Vector: PMD42nv; Purified genomic DNA from M.  
musculus C57BL/6J (male) was obtained from the Jackson



musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g1147321419b1AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 11 a 4 c 2 g 7 t  
ORIGIN

Query Match 100.0%; Score 7; DB 13; Length 24;  
Best Local Similarity 100.0%; Pred. No. 4.2e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 tttagta 7  
|||||||  
Db 24 TTTGATTA 18

Search completed: December 15, 2001, 02:33:44  
Job time: 4984 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 03:19:53 Search time 2725.73 Seconds  
(without alignments)  
133.153 Million cell updates/sec

Title: US-09-380-826a-6  
Perfect score: 1 tgttgannnnnnnttgata 22  
Sequence: 1 tgttgannnnnnnttgata 22

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:\*  
1: gb\_da:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sy:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_om:\*  
20: em\_ov:\*  
21: em\_ov:\*  
22: em\_pat:\*  
23: em\_ph:\*  
24: em\_pl:\*  
25: em\_ro:\*  
26: em\_sts:\*  
27: em\_sy:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_htgo\_hum:\*  
31: em\_htgo\_inv:\*  
32: em\_htgo\_rod:\*  
33: em\_htg\_hum:\*  
34: em\_htg\_inv:\*  
35: em\_htg\_rod:\*  
36: em\_htg\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	14	63.6	239	9	HSSSG03
2	14	63.6	435	11	HUM4STS161
3	14	63.6	668	11	G50458
4	14	63.6	1015	14	SRSSRSV
5	14	63.6	1257	4	BTBJ9225
6	14	63.6	1257	9	HSBJ9229
7	14	63.6	1258	4	BTBJ9224
8	14	63.6	1276	4	BTBJ9223
9	14	63.6	1287	9	HSUS2077
10	14	63.6	1305	6	AXI48809
11	14	63.6	1306	6	AXI48806
12	14	63.6	1481	1	LFU60594
13	14	63.6	2321	3	DDU66913
14	14	63.6	2683	1	SPGRDELN
15	14	63.6	2683	6	AXI48805
16	14	63.6	3666	6	A45819
17	14	63.6	3666	6	A45820
18	14	63.6	3915	8	SCYJL046W
19	14	63.6	4197	6	A45813
20	14	63.6	4197	6	A45814
21	14	63.6	4684	8	SCYJL045W
22	14	63.6	5620	10	MTTHREC02
23	14	63.6	7697	14	AB042808
24	14	63.6	10389	1	AE006627
25	14	63.6	10948	2	AC014767
26	14	63.6	11949	9	AE000145
27	14	63.6	12437	1	AE002102
28	14	63.6	12973	2	AC013026
29	14	63.6	12978	3	AC006805
30	14	63.6	26923	3	CEM0386
31	14	63.6	29000	3	AY027893S6
32	14	63.6	37923	3	AF047659
33	14	63.6	42042	3	CEC3484
34	14	63.6	59012	9	HSABLGR2
35	14	63.6	59909	2	AC090996
36	14	63.6	63156	2	AC026482
37	14	63.6	63659	2	AC090590
38	14	63.6	63739	9	AF271897
39	14	63.6	63933	2	AC090570
40	14	63.6	64577	3	AE002828
41	14	63.6	64789	2	AC083839
42	14	63.6	67987	2	AC016064
43	14	63.6	68082	2	AC014851
44	14	63.6	68206	2	AC087450
45	14	63.6	69358	2	AC090544

## ALIGNMENTS

RESULT 1  
LOCUS HSSSG03 239 bp DNA PRI 14-MAR-1999  
DEFINITION Homo sapiens squalene synthase gene, intron 2, 5' end.  
ACCESSION AF051649  
VERSION AF051649.1 GI:4415976

KEYWORDS  
SEGMENT  
SOURCE  
ORGANISM

Homo sapiens  
human.  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 239)  
Jiang, G., McKenzie, T.L., Conrad, D.G. and Shechter, I.  
Transcriptional regulation by lovastatin and 25-hydroxycholesterol  
in HepG2 cells and molecular cloning and expression of the cDNA for  
the human hepatic squalene synthase

JOURNAL  
MEDLINE  
REFERENCE  
J. Biol. Chem. 268 (117), 12818-12824 (1993)  
93286128  
2 (bases 1 to 239)

AUTHORS Guan, G., Dai, P.H. and Shechter, I.  
TITLE Direct Submission  
JOURNAL Submitted (02-MAR-1998) Biochemistry and Molecular Biology,  
Uniformed Services University of the Health Sciences, 4301 Jones  
Bridge Road, Bethesda, MD 20814-4799, USA  
FEATURES Location/Qualifiers  
source 1. 239  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="8"  
/map="8p22-p23.1"  
33. .>239  
/gene="squalene synthase"  
/number=2  
Intron  
BASE COUNT 43 a 61 c 70 g 65 t  
ORIGIN  
Query Match 63.6%; Score 14; DB 9; Length 239;  
Best Local Similarity 63.6%; Pred. No. 1.2e+03;  
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
Oy 1 tgttgagannnnnnnttgata 22  
|||||  
Db 92 TGTTCGAGCTACTCTTGATA 113  
RESULT 2  
HUM4STS161 435 bp DNA STS 26-FEB-1996  
LOCUS Hum. chromosome 4 sequence-tagged site STS4-161, sequence tagged  
DEFINITION site.  
ACCESSION M95263.1 GI:177262  
VERSION  
KEYWORDS STS; human chromosome 4; sequence tagged site.  
SOURCE Homo sapiens, clone C4-177 from Los Alamos National Laboratory  
chromosome 4 cosmid library, plate 4-1-2R, pos3c; vector SuperCos 1  
(StrataGene).  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 435)  
AUTHORS Gould, R.D., disibio, G., Xu, H., Lang, D.B., Dadgar, J., Magrane, G.,  
Dugalczyk, A., Smith, K.A., Cox, D.R., Masters, S.B. and Myers, R.M.  
TITLE The development of sequence-tagged sites for human chromosome 4  
JOURNAL Hum. Mol. Genet. 2 (8), 1271-1288 (1993)  
MEDLINE 94004872  
COMMENT PCR components: 25 ng of human genomic DNA, 10 pmol of each  
oligonucleotide, 200 micro-M dNTPs, 0.25 U Taq polymerase (cetus)  
in 10 micro-1 of 50 mM KCl-20 mM Tris-HCl, pH 8.3 (at room temp),  
2.5 mM MgCl-2. Initial denaturation at 94degC for 1.5 min, then 30  
cycles of 94degC for 15 sec, 62degC for 23 sec, and 72degC for 30  
sec, followed by a final extension at 72degC for 3.5 min, using a  
Perkin-Elmer 9600 thermocycler. PCR-amplified product size 279 bp.  
Sequence submitted by:  
Human Genome Mapping Center  
Box 0925  
University of California San Francisco  
San Francisco, CA 94143-0925 USA  
Phone: (415) 502-1612 Fax: (415) 476-8391  
e-mail: hgmprobe@cgl.ucsf.edu.  
FEATURES Location/Qualifiers  
source 1. 435  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
primer\_bind 316. .340  
primer\_bind 128 a 64 c 75 g 162 t 6 others  
BASE COUNT 128 a 64 c 75 g 162 t 6 others  
ORIGIN  
Query Match 63.6%; Score 14; DB 11; Length 435;  
Best Local Similarity 63.6%; Pred. No. 1.2e+03;

Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
Oy 1 tgttgagannnnnnnttgata 22  
|||||  
Db 262 TGTTCGAGCTACTCTTGATA 283  
RESULT 3  
G50458 668 bp DNA STS 30-MAR-2000  
LOCUS SHGC-79916 Human Homo sapiens STS genomic, sequence tagged site.  
DEFINITION  
ACCESSION G50458  
VERSION G50458.1 GI:5221635  
KEYWORDS STS.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 668)  
AUTHORS Olivier, M. and Cox, D.R.  
TITLE Unpublished, Olivier, M., Cox, D.R. (2000)  
JOURNAL Unpublished (2000)  
COMMENT  
Contact: Michael Olivier, David R. Cox  
Stanford Human Genome Center  
Stanford University, School of Medicine  
4005 Miranda Ave. 2nd Fl., Palo Alto, CA 94025, USA  
Tel: (650) 320-5800  
Fax: (650) 320-5801  
Email: olivier@shgc.stanford.edu  
Primer A: GCACATTAGCTTGGATTGCTT  
Primer B: TGCCCTTGCTTGGATTGATA  
STS size: 278  
PCR Profile:  
Initial incubation: 95 degrees C for 10 minutes  
Denaturation: 94 degrees C for 30 seconds  
Annealing: 60 degrees C for 30 seconds  
Polymerization: 72 degrees C for 23 seconds  
PCR Cycles: 30  
Thermal Cycler: Perkin Elmer 9700  
Protocol:  
Template: 25 ng  
Primer: each 1 uM  
dNTPs: each 200 uM  
AmpliTaq Gold Polymerase: 0.07 units/uL  
Total Vol: 5 uL  
Buffer:  
MgCl2: 2.5 mM  
KCl: 50 mM  
Tris-HCl: 10 mM  
pH: 8.3  
BAC ends sequenced at TIGR from the RPC11 BAC library. Designed  
and developed at the Stanford Human Genome Center.  
FEATURES Location/Qualifiers  
source 1. 668  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="8"  
/clone\_lib="Human"  
STS 31. .308  
primer\_bind 31. .53  
primer\_bind 230 a 104 c 113 g 221 t  
BASE COUNT 230 a 104 c 113 g 221 t  
ORIGIN  
Query Match 63.6%; Score 14; DB 11; Length 668;  
Best Local Similarity 63.6%; Pred. No. 1.2e+03;  
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
Oy 1 tgttgagannnnnnnttgata 22

Db 543 TGTGGAAATTAACCTTGATA 564

RESULT 4  
SRSSRSV/c 1015 bp RNA VRL 08-FEB-1999  
LOCUS Small round structured virus genomic RNA, 3' terminal sequence  
DEFINITION containing ORF2 and ORF3.  
ACCESSION D38547  
VERSION D38547.1 GI:560054  
KEYWORDS

ORGANISM  
Small round structured virus (isolate: patient; SRSV-CHIBA-407/87/J)  
cDNA to genomic RNA.  
small round structured virus  
viruses; ssRNA positive-strand viruses, no DNA stage;  
Caliciviridae; Norwalk-like viruses.

REFERENCE  
1 (bases 1 to 1015)  
Utagawa, E.T.  
Direct Submission  
Submitted (17-OCT-1994) to the DDBJ/EMBL/GenBank databases. Etsuko  
Shinju-Ku, Tokyo 162, Japan (tel:03-5285-1111(ex.2562),  
fax:03-5285-1177)  
2 (bases 1 to 1015)  
Utagawa, E.T., Takeda, N., Inouye, S., Kasuga, K. and Yamazaki, S.  
3'-terminal sequence of a small round structured virus (SRSV) in  
Japan  
Arch. Virol. 135 (1-2), 185-192 (1994)  
94256839

JOURNAL  
MEDLINE  
FEATURES  
source  
1. 1015  
/organism="small round structured virus"  
/isolate="patient; SRSV-CHIBA-407/87/J"  
/db\_xref="taxon:37141"  
/note="SRSVs are currently recognised as a potential human  
calicivirus."  
<1. 298  
/note="ORF2"  
/codon\_start=2  
/protein\_id="BA07549.1"  
/db\_xref="GI:1232112"  
/translation="VPCLLPEYITHTFISEQAPIQGEAALLHYVDPDTRNLGEFKLY  
PGYLTCVNPSSSTGPQPLPLDGVFVASFVSWRFYQLKPVGTAGPARGLGVR"  
298. 924  
/note="ORF3"  
/codon\_start=1  
/protein\_id="BA07550.1"  
/db\_xref="GI:1232113"  
/translation="MAQAIIICAIASASAGSLGAGIQAGAEALQAOQRYQODLTLOON  
SFNHDKEMLGYOMESNNKLAKNLNTRYSLLQAGLSSDAAAVAGAPVTRLYDMG  
VRVAPOSATTLRSNGNMAVPLPAOPKOKLASEGYSNPYADVORARVAVQSQNSS  
RMSVPRHQAQTYWTPPSTSSSVSTARGRFNDRLPLFANLR"

CDS  
/note="ORF2"  
/codon\_start=2  
/protein\_id="BA07549.1"  
/db\_xref="GI:1232112"  
/translation="VPCLLPEYITHTFISEQAPIQGEAALLHYVDPDTRNLGEFKLY  
PGYLTCVNPSSSTGPQPLPLDGVFVASFVSWRFYQLKPVGTAGPARGLGVR"  
298. 924  
/note="ORF3"  
/codon\_start=1  
/protein\_id="BA07550.1"  
/db\_xref="GI:1232113"  
/translation="MAQAIIICAIASASAGSLGAGIQAGAEALQAOQRYQODLTLOON  
SFNHDKEMLGYOMESNNKLAKNLNTRYSLLQAGLSSDAAAVAGAPVTRLYDMG  
VRVAPOSATTLRSNGNMAVPLPAOPKOKLASEGYSNPYADVORARVAVQSQNSS  
RMSVPRHQAQTYWTPPSTSSSVSTARGRFNDRLPLFANLR"

BASE COUNT 229 a 251 c 233 g 302 t  
ORIGIN

Query Match 63.6%; Score 14; DB 14; Length 1015;  
Best Local Similarity 63.6%; Pred. No. 1.1e+03;  
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Oy 1 tgttgannnnnnnttgata 22  
||||| |||||  
Db 481 TGTGGACATTTCATTGATA 460

RESULT 5  
BTAJ9225 1257 bp DNA MAM 05-JAN-2001  
LOCUS Bos taurus mariner related transposon Hsmar1, clone btmlec4.  
DEFINITION AJ009225  
ACCESSION AJ009225.1 GI:9187451  
KEYWORDS transposon.

SOURCE  
ORGANISM  
Bos taurus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
Bovidae; Bovinae; Bos.  
REFERENCE  
1 (bases 1 to 1257)  
Dematei, M.V., Auge-Gouillou, C., Pollet, N., Hamelin, M.H.,  
Meunier-Rotival, M. and Bigot, Y.  
Features of the mammal marl transposons in the human, sheep, cow,  
and mouse genomes and implications for their evolution  
Mamm. Genome 11 (12), 1111-1116 (2000)

JOURNAL  
MEDLINE  
REFERENCE  
2 (bases 1 to 1257)  
Bigot, Y.  
Direct Submission  
Submitted (04-JUN-1998) Bigot Y., Institut de Recherche sur la  
Biologie de l'Insecte, UPRESA CNRS 6035, Facult des Sciences, Parc  
Grandmont, 37200 Tours, FRANCE  
Location/Qualifiers  
1. 1257  
/organism="Bos taurus"  
/transposon="Hsmar1"  
/db\_xref="taxon:9913"  
/clone="btmlec4"

FEATURES  
source  
1. 1257  
/organism="Bos taurus"  
/transposon="Hsmar1"  
/db\_xref="taxon:9606"  
/clone="tmhsc3"

BASE COUNT 381 a 263 c 264 g 343 t  
ORIGIN

Query Match 63.6%; Score 14; DB 9; Length 1257;  
Best Local Similarity 63.6%; Pred. No. 1.1e+03;  
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Oy 1 tgttgannnnnnnttgata 22  
||||| |||||  
Db 36 TGTGGATTTCATTGATA 57

RESULT 6  
HSAJ9229 1257 bp DNA PRI 06-JAN-2001  
LOCUS Homo sapiens mariner related transposon Hsmar1, clone tmhsc3.  
DEFINITION AJ009229  
ACCESSION AJ009229.1 GI:9187522  
KEYWORDS transposon.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 1257)  
Dematei, M.V., Auge-Gouillou, C., Pollet, N., Hamelin, M.H.,  
Meunier-Rotival, M. and Bigot, Y.  
Features of the mammal marl transposons in the human, sheep, cow,  
and mouse genomes and implications for their evolution  
Mamm. Genome 11 (12), 1111-1116 (2000)

JOURNAL  
MEDLINE  
REFERENCE  
2 (bases 1 to 1257)  
Bigot, Y.  
Direct Submission  
Submitted (04-JUN-1998) Bigot Y., Institut de Recherche sur la  
Biologie de l'Insecte, UPRESA CNRS 6035, Facult des Sciences, Parc  
Grandmont, 37200 Tours, FRANCE  
Location/Qualifiers  
1. 1257  
/organism="Homo sapiens"  
/transposon="Hsmar1"  
/db\_xref="taxon:9606"  
/clone="tmhsc3"

BASE COUNT 381 a 263 c 264 g 343 t  
ORIGIN

Query Match 63.6%; Score 14; DB 9; Length 1257;  
Best Local Similarity 63.6%; Pred. No. 1.1e+03;  
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Matches 14: Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 tgttggannnnnnnnttgata 22  
|||||  
Db 36 TGTTCGAATTGGCATTGTGATA 57

RESULT 7  
BTAJ9224 1258 bp DNA MAM 05-JAN-2001  
LOCUS Bos taurus mariner related transposon Hsmar1, clone btmlec3.  
ACCESSION AJ009224  
VERSION AJ009224.1 GI:9187450  
KEYWORDS transposon.  
SOURCE cow.  
ORGANISM Bos taurus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
Bovidae; Bovinae; Bos.  
1 (bases 1 to 1258)  
Demattei,M.V., Auge-Gouillou,C., Pollet,N., Hamelin,M.H.,  
Meunier-Rotival,M. and Bigot,Y.  
Features of the mammal mar1 transposons in the human, sheep, cow,  
and mouse genomes and implications for their evolution  
Mamm. Genome 11 (12), 1111-1116 (2000)  
21015409  
Bigot,Y.  
Direct Submission  
Submitted (04-JUN-1998) Bigot Y., Institut de Recherche sur la  
Biologie de l'Insecte, UPRESA CNRS 6035, Facult des Sciences, Parc  
Grandmont, 37200 Tours, FRANCE  
Location/Qualifiers  
1. .1258  
/organism="Bos taurus"  
/transposon="Hsmar1"  
/db\_xref="taxon:9913"  
/clone="btmlec3"

BASE COUNT 386 a 262 c 266 g 344 t

ORIGIN

Query Match 63.6%; Score 14; DB 4; Length 1258;  
Best Local Similarity 63.6%; Pred. No. 1.1e+03;  
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 tgttggannnnnnnnttgata 22  
|||||  
Db 36 TGTTCGAATTGGCATTGTGATA 57

RESULT 8  
BTAJ9223 1276 bp DNA MAM 05-JAN-2001  
LOCUS Bos taurus mariner related transposon Hsmar1, clone btmlec2.  
ACCESSION AJ009223  
VERSION AJ009223.1 GI:9187449  
KEYWORDS transposon.  
SOURCE cow.  
ORGANISM Bos taurus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
Bovidae; Bovinae; Bos.  
1 (bases 1 to 1276)  
Demattei,M.V., Auge-Gouillou,C., Pollet,N., Hamelin,M.H.,  
Meunier-Rotival,M. and Bigot,Y.  
Features of the mammal mar1 transposons in the human, sheep, cow,  
and mouse genomes and implications for their evolution  
Mamm. Genome 11 (12), 1111-1116 (2000)  
21015409  
Bigot,Y.  
Direct Submission

JOURNAL Submitted (04-JUN-1998) Bigot Y., Institut de Recherche sur la  
Biologie de l'Insecte, UPRESA CNRS 6035, Facult des Sciences, Parc  
Grandmont, 37200 Tours, FRANCE  
Location/Qualifiers  
1. .1276  
/organism="Bos taurus"  
/transposon="Hsmar1"  
/db\_xref="taxon:9913"  
/clone="btmlec2"

BASE COUNT 388 a 261 c 248 g 379 t

ORIGIN

Query Match 63.6%; Score 14; DB 4; Length 1276;  
Best Local Similarity 63.6%; Pred. No. 1.1e+03;  
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 tgttggannnnnnnnttgata 22  
|||||  
Db 37 TGTTCGAATTGGCATTGTGATA 58

RESULT 9  
HSU52077 1287 bp DNA PRI 26-JAN-1998  
LOCUS Human mariner1 transposase gene, complete consensus sequence.  
DEFINITION U52077  
ACCESSION U52077  
VERSION U52077.1 GI:1263080  
KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
1 (bases 1 to 1287)  
Robertson,H.M. and Zampano,K.L.  
The human genome  
Molecular evolution of an ancient mariner transposon, Hsmar1, in  
Gene 205 (1-2), 203-217 (1997)  
98121293  
2 (bases 1 to 1287)  
Robertson,H.M.  
Direct Submission  
Submitted (21-MAR-1996) Hugh M. Robertson, Entomology, University  
of Illinois at Urbana-Champaign, 505 S. Goodwin, Urbana, IL 61801,  
USA  
Location/Qualifiers  
1. .1287  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/transposon="Hsmar1"  
/note="consensus sequence based on 20 unique sequences"

FEATURES  
source  
1. .30  
/rpt\_type=Inverted  
repeat\_region  
179..1210  
/codon\_start=-1  
/product="mariner transposase"  
/protein\_id="AAC52010.1"  
/db\_xref="GI:1263081"  
/translation="MEMMLDKQIRAIIFEEFKMKRAETTRNNINNAAGPGTANERT  
VQWFFKFKGDESLDEERSGRPSEVDNQLAIIEADPLTTREVAELNDHSTV  
VRHKKOIGKVKLDKRWPHLSQNKRRREVSLSLRNNNEPFLDRIVTDEKIL  
YDNRPRPAQWLDEEAPKPKPKLNOKKRVVYVMSAAGLIHSFLNNGEITTSKRY  
AQQIDENHRLQLQPALVNRKGPILLHDNARPHVAOPTLOKLNELGYEVLPHPPYSP  
DLSPDTDFEFKHLDNFLQGRFRHNOQDAENAFQEPYBSRSTDTYATGINKLISRWQK  
VDSNGSYFD"

repeat\_region  
1256..1287  
/rpt\_type=Inverted  
BASE COUNT 381 a 284 c 278 g 344 t

ORIGIN

Query Match 63.6%; Score 14; DB 9; Length 1287;  
Best Local Similarity 63.6%; Pred. No. 1.1e+03;

Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 ttttgagannnnnnnttgata 22  
|||||

Db 36 TGTGTGAGATTTCCTACTTGATA 57

RESULT 10  
AX148809/c  
LOCUS AX148809 1305 bp DNA  
DEFINITION Sequence 11 from Patent WO0136625.  
ACCESSION AX148809  
VERSION AX148809.1 GI:14347333  
KEYWORDS  
SOURCE Streptococcus pyogenes.  
ORGANISM Streptococcus pyogenes.  
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
Streptococcus.

REFERENCE  
AUTHORS Wright,J.A., Young,A.H. and Dugourd,D.  
TITLE Antisense oligonucleotide sequences derived from groEL and groES as  
inhibitors of microorganisms  
JOURNAL Patent: WO 0136625-A 11 25-MAY-2001;  
Genesense Technologies Inc. (CA)  
FEATURES  
source 1. .1305  
/organism="Streptococcus pyogenes"  
/db\_xref="taxon:1314"

BASE COUNT 406 a 238 c 312 g 349 t  
ORIGIN

Query Match 63.6%; Score 14; DB 6; Length 1305;  
Best Local Similarity 63.6%; Pred. No. 1.1e+03;  
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 ttttgagannnnnnnttgata 22  
|||||

Db 661 TGTGTGAGCTCCTACTTGATA 640

RESULT 11  
AX148806/c  
LOCUS AX148806 1306 bp DNA  
DEFINITION Sequence 8 from Patent WO0136625.  
ACCESSION AX148806  
VERSION AX148806.1 GI:14347330  
KEYWORDS  
SOURCE Streptococcus pyogenes.  
ORGANISM Streptococcus pyogenes.  
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
Streptococcus.

REFERENCE  
AUTHORS Wright,J.A., Young,A.H. and Dugourd,D.  
TITLE Antisense oligonucleotide sequences derived from groEL and groES as  
inhibitors of microorganisms  
JOURNAL Patent: WO 0136625-A 8 25-MAY-2001;  
Genesense Technologies Inc. (CA)  
FEATURES  
source 1. .1306  
/organism="Streptococcus pyogenes"  
/db\_xref="taxon:1314"

BASE COUNT 407 a 235 c 308 g 356 t  
ORIGIN

Query Match 63.6%; Score 14; DB 6; Length 1306;  
Best Local Similarity 63.6%; Pred. No. 1.1e+03;  
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 ttttgagannnnnnnttgata 22  
|||||

Db 667 TGTGTGAGCTCCTACTTGATA 646

RESULT 12  
LFU60594  
LOCUS LFU60594 1481 bp DNA BCT 10-SEP-1998  
DEFINITION Leptospiira fainel 16S ribosomal RNA gene, partial sequence.  
ACCESSION U60594  
VERSION U60594.1 GI:1408219  
KEYWORDS  
SOURCE Leptospiira fainel.  
ORGANISM Leptospiira fainel.  
Bacteria; Spirochaetales; Leptospiraceae; Leptospira.

REFERENCE  
AUTHORS Perolat,P., Chappel,R.J., Adler,B., Baranton,G., Bulach,D.M.,  
Billinghurst,M.L., Letcart,M., Merlen,F. and Serrano,M.S.  
TITLE Leptospiira fainel sp. nov., isolated from pigs in Australia  
JOURNAL Int. J. Syst. Bacteriol. 48 Pt 3, 851-858 (1998)  
MEDLINE 98404550  
REFERENCE 2 (bases 1 to 1481)  
AUTHORS Adler,B., Chappel,R.J., Baranton,G., Bulach,D.M.,  
Billinghurst,M.L., Letcart,M., Merlen,F., Serrano,M.S. and  
Perolat,P.  
TITLE Direct Submission  
JOURNAL Submitted (12-JUN-1996) Microbiology, Monash University, Wellington  
Rd., Clayton, VIC 3168, Australia  
FEATURES  
source 1. .1481  
/organism="Leptospiira fainel"  
/strain="Hurstbridge"  
/db\_xref="taxon:48782"

BASE COUNT 391 a 335 c 439 g 314 t 2 others  
ORIGIN

Query Match 63.6%; Score 14; DB 1; Length 1481;  
Best Local Similarity 63.6%; Pred. No. 1.1e+03;  
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 ttttgagannnnnnnttgata 22  
|||||

Db 155 TGTGTGATCACAAGATTTGATA 176

RESULT 13  
DDU66913/c  
LOCUS DDU66913 2321 bp DNA INV 05-SEP-1996  
DEFINITION Dictyostellium discoideum ORF DG1040 gene, partial cds.  
ACCESSION U66913  
VERSION U66913.1 GI:1519537  
KEYWORDS  
SOURCE Dictyostellium discoideum.  
ORGANISM Dictyostellium discoideum.  
Eukaryota; Mycetozoa; Dictyostellida; Dictyostellium.  
REFERENCE  
AUTHORS Loomis,W.F.  
TITLE Direct Submission  
JOURNAL Submitted (15-AUG-1996) Dept. of Biology 0322, University of  
California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA  
FEATURES  
source 1. .2321  
/organism="Dictyostellium discoideum"  
/strain="AX4"  
/db\_xref="taxon:44689"  
/cell\_line="AX4"  
join(370..732,813..>2321)  
/note="ORF DG1040"  
<370..732  
/number=1  
join(370..732,813..>2321)  
/note="ORF DG1040; Initially derived from a plasmid  
disrupted gene"

```
/codon_start=1
/protein_id="AAB07544.1"
/db_xref="GI:1519538"
/translation="MSTRPESEDLKSSSEVNDSTINNONNNNSINNNNDNNNNKKEE
TSEYDPPNNTEFNKSLSSDKFYKFWLQVSSNNNEQTSNTISNLSVATIKN
LISNLTJKSKYSKNSSELQIINQVGOAEIYIFRCLVDSIDKRNTPITPTTSPTTSA
TSTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPTPT
VNNNEFLVFSSTLKTPTSHETIMIGLALSOSIDKRVREDAKFLINTLITOLSNTKS
LPENLHOLLYKFKTOETNOLOOONNOOONNOOONNOOONNOOONNOOONNOOONNO
OOONOOHSEVTKYTILOSISPPNSALSPLYNDOSGSGGGSITSKRIQDNPSYVN
SIISKVSQONVPYQIILSISYSSCSASTLKILFQFKLTESDIAQIILVMAELTSN
TNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NSNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
LAVYKATQVDDQSKPPIDFLIDRWVWNSLSQIOLKIAIOSDFEPFQNSTKKRID"
exon
813..>2321
/number=2
misc_feature
2124
/note="site of plasmid insertion in mutant organisms;
mutants form crinkled, branching fingers; neither spores
nor stalk cells were seen"
BASE COUNT      1055 a      338 c      210 g      718 t
ORIGIN
Query Match      63.6%; Score 14; DB 3; Length 2321;
Best Local Similarity 63.6%; Pred. No. 1.1e+03;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY      1 tgttgannnnnnnttgata 22
|||||
Db      2069 TGTGGATATTAATCTTGTGATA 2048

RESULT 14
SPGROELGN 2683 bp DNA BCT 26-SEP-1997
LOCUS
DEFINITION
S.pyogenes DNA for groEL gene.
X89236
X89236.1 GI:2462691
groEL gene: heat shock protein 60 (GroEL) like protein.
SOURCE
Streptococcus pyogenes.
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.
REFERENCE
1 (bases 1 to 2683)
Podbielski, A.
Direct Submission
Submitted (29-JUN-1995) A. Podbielski, Institut f Med
Mikrobiologie, an der Rhein Westf. Techn. Hochschule, Aachen,
Pauwelsstr Klinikum, 52057 Aachen, FRG
2 (bases 1 to 2683)
Pohl, B., Podbielski, A. and Zarges, I.
Unpublished
Related sequences M81132, M84965.
FEATURES
Location/Qualifiers
1..2683
/organism="Streptococcus pyogenes"
/strain="serotype M49"
/isolate="CS101"
/db_xref="taxon:1314"
1..1446
/gene="groEL"
<1..1446
/gene="groEL"
/codon_start=1
/transl_table=11
/product="heat shock protein"
/protein_id="CA61520.1"
/db_xref="GI:2462692"
/db_xref="SPRMBL:033733"
/translation="DHFEWNGAKLVSEVASKTNDIAGDGTATATVITQALVHGELKNV
TAGANPIGIRGILETATATAVALKAVAPVSGKEAIAOVAASRSREKVGEXTISEAM
```

```
BRVNDGVITTEESRGMETELEVVEGQFDRGYLSQYVWTDNEKNVADLENPFILMD
KRVSNIDILPLEEVLKTNRPLLIADVDGEALPTVLNKRGTFFNVAVKAPGFG
DRKRAMLEDIALVGGVLTEDLGELEADATWATLGOAKITVDSDTVYVSGSGSE
AIANRILIKSOLETTTSDPDEKIOEERAKIAGGAVYIKVCAPEETLKEKKLIED
AINATRAVEESIVAGGATLITTEKVALEBEDDAGRIVIRALEEYRQALN
AGEESVVIDKLKNSPAGFNAATGEVWDIMKTGIIIDPVKTRSAIONAASVALIL
TTEAVVANKPEPAPAPAMPACMDGMMGGF"
repeat_region
1509..1564
/rpt_type=INVERTED
BASE COUNT      808 a      490 c      593 g      792 t
ORIGIN
Query Match      63.6%; Score 14; DB 1; Length 2683;
Best Local Similarity 63.6%; Pred. No. 1.1e+03;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY      1 tgttgannnnnnnttgata 22
|||||
Db      963 TGTGGAGCTCCTACTTGTGATA 942

RESULT 15
AXI48805 2683 bp DNA PAT 08-JUN-2001
LOCUS
DEFINITION
Sequence 7 from Patent WO0136625.
AXI48805
VERSION
AXI48805.1 GI:14347329
KEYWORDS
Streptococcus pyogenes.
Streptococcus pyogenes
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.
REFERENCE
1 (bases 1 to 2683)
Wright, J.A., Young, A.H. and Dugourd, D.
Antisense oligonucleotide sequences derived from groEL and groES as
inhibitors of microorganisms
Patent: WO 0136625-A 7 25-MAY-2001;
GeneSense Technologies Inc. (CA)
FEATURES
Location/Qualifiers
1..2683
/organism="Streptococcus pyogenes"
/db_xref="taxon:1314"
BASE COUNT      808 a      490 c      593 g      792 t
ORIGIN
Query Match      63.6%; Score 14; DB 6; Length 2683;
Best Local Similarity 63.6%; Pred. No. 1.1e+03;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY      1 tgttgannnnnnnttgata 22
|||||
Db      963 TGTGGAGCTCCTACTTGTGATA 942
```

Search completed: December 15, 2001, 03:19:57  
Job time: 7687 sec





**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 03:29:44 : Search time 401.91 Seconds  
(without alignments)  
46.922 Million cell updates/sec

Title: US-09-380-826a-6  
Perfect score: 22  
Sequence: 1 tglgtgannnnnnnttgcata 22

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: /SID52/gcgdata/geneseq/NA1980.DAT:\*  
2: /SID52/gcgdata/geneseq/NA1981.DAT:\*  
3: /SID52/gcgdata/geneseq/NA1982.DAT:\*  
4: /SID52/gcgdata/geneseq/NA1983.DAT:\*  
5: /SID52/gcgdata/geneseq/NA1984.DAT:\*  
6: /SID52/gcgdata/geneseq/NA1985.DAT:\*  
7: /SID52/gcgdata/geneseq/NA1986.DAT:\*  
8: /SID52/gcgdata/geneseq/NA1987.DAT:\*  
9: /SID52/gcgdata/geneseq/NA1988.DAT:\*  
10: /SID52/gcgdata/geneseq/NA1989.DAT:\*  
11: /SID52/gcgdata/geneseq/NA1990.DAT:\*  
12: /SID52/gcgdata/geneseq/NA1991.DAT:\*  
13: /SID52/gcgdata/geneseq/NA1992.DAT:\*  
14: /SID52/gcgdata/geneseq/NA1993.DAT:\*  
15: /SID52/gcgdata/geneseq/NA1994.DAT:\*  
16: /SID52/gcgdata/geneseq/NA1995.DAT:\*  
17: /SID52/gcgdata/geneseq/NA1996.DAT:\*  
18: /SID52/gcgdata/geneseq/NA1997.DAT:\*  
19: /SID52/gcgdata/geneseq/NA1998.DAT:\*  
20: /SID52/gcgdata/geneseq/NA1999.DAT:\*  
21: /SID52/gcgdata/geneseq/NA2000.DAT:\*  
22: /SID52/gcgdata/geneseq/NA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	14	63.6	22	19	AAV58901	Leptospira RNA ge
2	14	63.6	22	19	AAV58897	L. finnel nucleoti
3	14	63.6	22	19	AAV58900	Leptospira nucleot
4	14	63.6	500	22	AAI15869	Probe #5802 for ge
5	14	63.6	500	22	AAI37774	Probe #6460 used t
6	14	63.6	972	20	AAK39627	Nucleic acid seque
c 7	14	63.6	1305	22	AAH56865	Antibiotic resista
c 8	14	63.6	1306	22	AAH56862	Antibiotic resista
9	14	63.6	1477	19	AAV58896	L. finnel nucleoti
10	14	63.6	1573	22	AAH47810	Human mariner tran
c 11	14	63.6	1661	20	AAH86155	Human encoding a Str

12	14	63.6	2660	21	AAK47974	Arabidopsis thalia
c 13	14	63.6	2683	22	AAH56861	S. pyogenes groEL
c 14	14	63.6	4197	16	AAO99430	B. sphaericus SLP
c 15	13	59.1	149	21	AAC28622	Human secreted pro
c 16	13	59.1	278	22	AAI71057	C. glutamicum SRT
17	13	59.1	396	20	AAK34752	DNA encoding OMP-1
18	13	59.1	430	21	AAK09647	Human secreted pro
19	13	59.1	664	22	AAK22692	Human gastric canc
20	13	59.1	828	20	AAI26735	Human gene express
21	13	59.1	1032	22	AAK23184	DNA encoding novel
c 22	13	59.1	1079	20	AAK20676	Polynucleotide seq
c 23	13	59.1	1092	22	AAK23271	DNA encoding novel
c 24	13	59.1	1266	22	AAK88112	Human FLEXHT-43 nu
c 25	13	59.1	1482	19	AAV18099	Nucleotide sequenc
c 26	13	59.1	1669	22	AAK90031	SPO ID 1. Staphyl
c 27	13	59.1	1786	18	AAI60350	MAP kinase #2 codi
c 28	13	59.1	2084	21	AAV79696	Eucalyptus grandis
c 29	13	59.1	2133	22	AAK11387	Human phospholipas
c 30	13	59.1	2190	20	AAK07102	Staphylococcus aur
c 31	13	59.1	2340	22	AAH66171	C. glutamicum codin
c 32	13	59.1	2463	22	AAI71056	C. glutamicum SRT
c 33	13	59.1	2508	18	AAI67197	Zebrafish retinoid
c 34	13	59.1	3487	22	AAK80522	Receptor #10 parti
c 35	13	59.1	3519	21	AAK294941	Human carbohydrate
c 36	13	59.1	3519	21	AAK294948	Human carbohydrate
c 37	13	59.1	4012	18	AAV74358	Staphylococcus aur
c 38	13	59.1	4254	22	AAI61065	Human polynucleoti
c 39	13	59.1	4685	22	AAI59279	Human polynucleoti
c 40	13	59.1	4803	22	AAI11574	Human phospholipas
c 41	13	59.1	4820	22	AAI11388	Human phospholipas
c 42	13	59.1	4875	22	AAI11386	Human phospholipas
c 43	13	59.1	5253	21	AAK26868	Essential Staphylo
c 44	13	59.1	5253	22	AAK08019	Staphylococcus aur
45	13	59.1	5253	22	AAI91550	Staphylococcus aur

ALIGNMENTS

RESULT 1	
AAV58901	
ID AAV58901 standard; DNA; 22 BP.	
XX AC AAV58901;	
XX DT 20-JAN-1999 (first entry)	
XX DE Leptospira RNA gene nucleotide sequence.	
XX DE Infection; pathogenic Leptospira; protective immunity; therapy;	
KW diagnosis; ss.	
KW OS Leptospira sp.	
XX OS	
XX PN W09840099-A1.	
XX PD 17-SEP-1998.	
XX PF 06-MAR-1998; 98WO-AU00145.	
XX PR 07-MAR-1997; 97AU-0005494.	
XX PA (AGRI-) AGRIC VICTORIA SERVICES PTY LTD.	
XX PA (PIGR-) PIG RES & DEV CORP.	
XX PI Chappel RJ;	
DR WPI; 1998-520791/44.	
XX PT New isolated pathogenic Leptospira bacterium - useful for, e.g	
PT developing products for conferring protective immunity, and for	
PT prophylactic or therapeutic treatment	
XX	

PS Claim 15; Page 72; 94pp; English.

XX This sequence represents a Leptospira DNA sequence isolated from the  
CC pathogenic Leptospira (LS) bacterium of the invention. The bacterium  
CC belongs to serogroup Hurstbridge or serovar hurstbridge or the species  
CC L. fainei. The LS bacteria can be used for conferring protective  
CC immunity against pathogenic LS bacteria in humans or animals. The  
CC bacteria can also be used for prophylactic or therapeutic treatment of LS  
CC infections. The DNAs and antibodies may also be used for detection and  
CC diagnosis of past or present LS infection.

XX Sequence 22 BP; 7 A; 2 C; 5 G; 8 T; 0 other;

SO

Query Match 63.6%; Score 14; DB 19; Length 22;  
Best Local Similarity 63.6%; Pred. No. 77;  
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 ttttgagannnnnnnttgata 22  
||||| |||||  
Db 1 ttttgatcacaagatttgata 22

RESULT 2

AAV58897  
ID AAV58897 standard; DNA; 22 BP.

AC AAV58897;  
XX  
XX 20-JAN-1999 (first entry)  
DT  
XX L. fainei nucleotide sequence.  
DE  
XX Infection; pathogenic Leptospira; protective immunity; therapy;  
XX diagnosis; ss.  
KW  
XX Leptospira fainei.  
OS  
XX WO9840099-A1.  
PN  
XX 17-SEP-1998.  
PD  
XX 06-MAR-1998; 98WO-AU00145.  
PF  
XX 07-MAR-1997; 97AU-0005494.  
PR  
XX (AGRI-) AGRIC VICTORIA SERVICES PTY LTD.  
PA (PIGR-) PIG RES & DEV CORP.  
XX Chappel RJ;  
PI  
XX WPI; 1998-520791/44.  
DR  
XX New isolated pathogenic Leptospira bacterium - useful for, e.g  
PT developing products for conferring protective immunity, and for  
PT prophylactic or therapeutic treatment  
XX  
XX Claim 15; Page 70; 94pp; English.

CC This sequence represents a Leptospira DNA sequence isolated from the  
CC pathogenic Leptospira (LS) bacterium of the invention. The bacterium  
CC belongs to serogroup Hurstbridge or serovar hurstbridge or the species  
CC L. fainei. The LS bacteria can be used for conferring protective  
CC immunity against pathogenic LS bacteria in humans or animals. The  
CC bacteria can also be used for prophylactic or therapeutic treatment of LS  
CC infections. The DNAs and antibodies may also be used for detection and  
CC diagnosis of past or present LS infection.

XX Sequence 22 BP; 7 A; 2 C; 5 G; 8 T; 0 other;

SO

Query Match 63.6%; Score 14; DB 19; Length 22;  
Best Local Similarity 63.6%; Pred. No. 77;  
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 ttttgagannnnnnnttgata 22  
||||| |||||  
Db 1 ttttgatcacaagatttgata 22

RESULT 3

AAV58900  
ID AAV58900 standard; DNA; 22 BP.

AC AAV58900;  
XX  
XX 20-JAN-1999 (first entry)  
DT  
XX Leptospira nucleotide sequence.  
DE  
XX Infection; pathogenic Leptospira; protective immunity; therapy;  
KW diagnosis; ss.  
XX  
XX Leptospira sp.  
OS  
XX WO9840099-A1.  
PN  
XX 17-SEP-1998.  
PD  
XX 06-MAR-1998; 98WO-AU00145.  
PF  
XX 07-MAR-1997; 97AU-0005494.  
PR  
XX (AGRI-) AGRIC VICTORIA SERVICES PTY LTD.  
PA (PIGR-) PIG RES & DEV CORP.  
XX Chappel RJ;  
PI  
XX WPI; 1998-520791/44.  
DR  
XX New isolated pathogenic Leptospira bacterium - useful for, e.g  
PT developing products for conferring protective immunity, and for  
PT prophylactic or therapeutic treatment  
XX  
XX Claim 15; Page 72; 94pp; English.

CC This sequence represents a Leptospira DNA sequence isolated from the  
CC pathogenic Leptospira (LS) bacterium of the invention. The bacterium  
CC belongs to serogroup Hurstbridge or serovar hurstbridge or the species  
CC L. fainei. The LS bacteria can be used for conferring protective  
CC immunity against pathogenic LS bacteria in humans or animals. The  
CC bacteria can also be used for prophylactic or therapeutic treatment of LS  
CC infections. The DNAs and antibodies may also be used for detection and  
CC diagnosis of past or present LS infection.

XX Sequence 22 BP; 3 A; 0 C; 4 G; 7 T; 8 other;

SO

Query Match 63.6%; Score 14; DB 19; Length 22;  
Best Local Similarity 100.0%; Pred. No. 77;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttttgagannnnnnnttgata 22  
||||| ||||| ||||| |||||  
Db 1 ttttgagannnnnnnttgata 22

RESULT 4

AAI15869  
ID AAI15869 standard; DNA; 500 BP.

AC AAI15869;  
XX  
XX 12-OCT-2001 (first entry)  
DT  
XX Probe #5802 for gene expression analysis in human cervical cell sample.  
DE

Pf	30-JAN-2001; 2001WO-US00663.
Px	
Pz	04-FEB-2000; 2000US-0180312.
PR	26-MAY-2000; 2000US-0207456.
PR	30-JUN-2000; 2000US-0608408.
PR	03-AUG-2000; 2000US-0632366.
PR	21-SEP-2000; 2000US-0234687.
PR	27-SEP-2000; 2000US-0236359.
PR	04-OCT-2000; 2000GB-0024263.
PX	
PA	(MOLE-) MOLECULAR DYNAMICS INC.
XX	
PI	Penn SG, Hanzel DK, Chen W, Rank DR;
XX	
DR	WPI; 2001-488897/53.
XX	
Pt	Human genome-derived single exon nucleic acid probes useful for
Pt	analyzing gene expression in human placenta -
XX	
PS	Claim 25; SEQ ID No 6460; 654bp; English.
XX	
CC	The present invention relates to single exon nucleic acid probes (SENP).
CC	The present sequence is one such probe. The probes are useful for
CC	producing a microarray for predicting, measuring and displaying gene
CC	expression in samples derived from human placenta. The probes are useful
CC	for antenatal diagnosis of human genetic disorders.
XX	
SQ	Sequence 500 BP; 171 A; 80 C; 92 G; 157 T; 0 other;
Oy	
Oy	Query Match                  63.6%; Score 14; DB 22; Length 500;
	Best Local Similarity        63.6%; Pred.No. 1.1e+02;
	Matches     14; Conservative     0; Mismatches     8; Indels     0; Gaps     0;
	1 tgttgannnnnnntttgata 22
Db	65 tgttgaatttcattcgata 86
RESULT	6
AAx99627/c	
ID	AAx99627 standard; DNA; 972 BP.
AC	AAx99627;
XX	
Dt	05-OCT-1999 (first entry)
XX	
DE	Nucleic acid sequence from U. urealyticum.
XX	
KW	Ureaplasma urealyticum; nucleic acid detection; infection; pathogen;
KW	human urogenital tract; pregnancy; neonatal disease; drug therapy;
KW	suppurative arthritis; ss.
XX	
OS	Ureaplasma urealyticum.
PN	
PN	WO9939007-A1.
XX	
PD	05-AUG-1999.
XX	
Pf	29-JAN-1999; 99WO-US01972.
XX	
PR	30-JAN-1998; 98US-0073189.
XX	
PA	(UABR-) UAB RES FOUND.
XX	
PI	Cassell GH, Chen EY, Glass JI, Glass JS, Heiner CR;
PI	Iefkowitz E;
XX	
DR	WPI; 1999-469343/39.
XX	
Pt	Detection of Ureaplasma urealyticum using novel genes, probes and
Pt	primers
XX	

PS Claim 1; Page 82; 110pp; English.

XX The present invention provides methods for the detection and diagnosis

CC of ureaplasma urealyticum infection. It provides novel genes

CC (AA99501-681) that can be used as a source of primers and probes for the

CC detection and/or quantification of U. urealyticum in a biological

CC sample. The probes that can be used in the method of the invention by

CC forming target:probe complex is complementary to a region selected from

CC one of the 181 nucleic acid sequences (AA99501-681). U. urealyticum is

CC an opportunistic pathogen of the human urogenital tract that is a

CC significant cause of adverse pregnancy outcome, neonatal disease, and

CC suppurative arthritis. As the infections are commonly asymptomatic, it is

CC important to have specific and sensitive methods for detecting their

CC presence in a patient. Also, as the pathogen has no current antibiotic

CC directed specifically against it, it would be advantageous to isolate and

CC detect gene sequences which are unique to it, and utilize these as a

CC basis for diagnosis of U. urealyticum infection as well as to develop new

CC and improved drug therapies. The present invention provides such novel

CC polynucleotide sequences (AA99501-681).

CC

XX Sequence 972 BP; 309 A; 127 C; 98 G; 438 T; 0 other;

XX

QY Query Match 63.6%; Score 14; DB 20; Length 972;

XX Best Local Similarity 63.6%; Pred. No. 1.2e+02;

XX Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

XX

QY 1 ttttgannnnnnnttgata 22

XX ||||||| |||||||

DB 931 TGTGGACTATTAGCTTGCATA 910

XX

RESULT 7

AAH56865/c

ID AAH56865 standard; DNA; 1305 BP.

XX

AC AAH56865;

XX

DT 06-SEP-2001 (first entry)

XX

DE Antibiotic resistant S. pyogenes groEL gene partial sequence SEQ ID-11.

XX

XX Antisense oligonucleotide; groE; groEL; groES; inhibitor; growth;

KW microorganism; Escherichia coli; Streptococcus pneumoniae; diagnosis;

KW Streptococcus pyogenes; Staphylococcus aureus; Pseudomonas aeruginosa;

KW antibacterial; antiviral; antiproliferative; antisense therapy;

KW microbial infection; ds.

XX

OS Streptococcus pyogenes.

XX

PN WO200136625-A2.

XX

PD 25-MAY-2001.

XX

PF 20-NOV-2000; 2000WO-CA01347.

XX

PR 18-NOV-1999; 99US-0166249.

XX

PA (GENE-) GENESENSE TECHNOLOGIES INC.

XX

PI Wright JA, Young AH, Dugourd D;

XX

DR WPI; 2001-355633/37.

XX

XX Novel antisense compounds targeting nucleic acid encoding groEL or

PT groES gene of microorganism, which hybridize with and inhibit

PT expression of the genes, useful to inhibit growth of microorganism

PT having the genes -

PS Disclosure; Fig 11; 110pp; English.

XX

XX The present invention specifically claims AAH56368 to AAH56832 which are

CC antisense oligonucleotides to nucleotide sequences encoding groE. More

CC generally, antisense compounds (I) comprising antisense oligonucleotides

CC of 5-50 bases targeted to a nucleotide sequence encoding groEL (heat

CC shock protein (HSP)60) (GL and groES (HSP10) (GS) gene from a

CC microorganism, where the antisense compound is complementary to GL or

CC GS of a microorganism and specifically hybridizes with and inhibits the

CC expression of GL or GS, is claimed. (I) have antibacterial, antiviral

CC and antiproliferative activities, and can be used in antisense therapy

CC and for inhibition of expression of groES or groEL. (I) are useful for

CC inhibiting expression of GL or GS in cells or tissues in vitro. (I) are

CC also useful for inhibiting the growth of a microorganism, or inhibiting

CC the expression of GL or GS gene in a microorganism (a bacterial cell or

CC a virus) having a GL or GS gene which involves administering to the

CC microorganism or to a cell infected with the microorganism, (I). (I) are

CC also useful for treating a mammalian pathological condition mediated by

CC the microorganisms which involves identifying a eukaryotic organism

CC having a pathological condition mediated by microorganisms having a GL

CC or GS gene and administering (I) such that the growth of microorganism

CC is inhibited. The antisense compounds are utilized for diagnostics,

CC therapeutics, prophylaxis and as research reagents and kits, e.g., to

CC prevent or delay microbial infections in humans. They are also useful as

CC molecular weight markers. AAH56362 to AAH56367 and AAH56833 to AAH56854

CC represent PCR primers for groE sequences which are used in the

CC exemplification of the present invention. AAH56855 to AAH56870 represent

CC groE nucleotide sequence given in the present invention.

XX

XX Sequence 1305 BP; 406 A; 238 C; 312 G; 349 T; 0 other;

XX

QY Query Match 63.6%; Score 14; DB 22; Length 1305;

XX Best Local Similarity 63.6%; Pred. No. 1.3e+02;

XX Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

XX

QY 1 ttttgannnnnnnttgata 22

XX ||||||| |||||||

DB 661 TGTGGAGCTCTACTTGCATA 640

XX

RESULT 8

AAH56862/c

ID AAH56862 standard; DNA; 1306 BP.

XX

AC AAH56862;

XX

DT 06-SEP-2001 (first entry)

XX

DE Antibiotic resistant S. pyogenes groEL gene partial sequence SEQ ID-8.

XX

XX Antisense oligonucleotide; groE; groEL; groES; inhibitor; growth;

KW microorganism; Escherichia coli; Streptococcus pneumoniae; diagnosis;

KW Streptococcus pyogenes; Staphylococcus aureus; Pseudomonas aeruginosa;

KW antibacterial; antiviral; antiproliferative; antisense therapy;

KW microbial infection; ds.

XX

OS Streptococcus pyogenes.

XX

PN WO200136625-A2.

XX

PD 25-MAY-2001.

XX

PF 20-NOV-2000; 2000WO-CA01347.

XX

PR 18-NOV-1999; 99US-0166249.

XX

PA (GENE-) GENESENSE TECHNOLOGIES INC.

XX

PI Wright JA, Young AH, Dugourd D;

XX

DR WPI; 2001-355633/37.

XX

XX Novel antisense compounds targeting nucleic acid encoding groEL or

PT groES gene of microorganism, which hybridize with and inhibit

PT expression of the genes, useful to inhibit growth of microorganism

PT having the genes -



CC Immunological diseases and inflammation.  
XX  
SQ Sequence 1573 BP; 538 A; 297 C; 304 G; 434 T; 0 other;

Query Match 63.6%; Score 14; DB 22; Length 1573;  
Best Local Similarity 63.6%; Pred. No. 1.3e+02;  
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 tgttgagannnnmttgata 22  
|||||||  
DB 415 tgttggaattgcatttgata 436

RESULT 11  
AAx86155/c  
ID AAx86155 standard; DNA; 1661 BP.  
XX  
AC AAx86155;  
XX  
DT 22-SEP-1999 (first entry)  
XX  
DE DNA encoding a Streptococcus pyogenes heat shock protein (Hsp)60-2.  
XX  
KW Heat shock protein; Hsp60-2; Immune response; immunological carrier;  
KW cancer control; tumour; sarcoma; cancer; gene therapy; ss.  
XX  
OS Streptococcus pyogenes.  
XX  
PN W09935270-A1.  
XX  
PD 15-JUL-1999.  
XX  
PF 29-DEC-1998; 98MO-CA01203.  
XX  
PR 31-DEC-1997; 97US-0001737.  
XX  
(STRE-) STRESSGEN BIOTECHNOLOGIES CORP.  
XX  
PI Mlzen L, Wisniewski J;  
XX  
DR WPI: 1999-430397/36.  
XX  
P-PSDB; MAY23904.  
XX  
PT New nucleic acid encoding heat shock protein-60 from Streptococcus,  
XX  
PT useful in vaccines, as carriers for other immunogens, as anticancer  
XX  
PT agents and for diagnosis  
XX  
PS Claim 3; Fig 4A-B; 176pp; English.  
XX  
XX  
CC The present sequence encodes a heat shock protein, designated Hsp60-2.  
CC The protein, its fragments, variants and fusion proteins, are  
CC used to elicit or enhance an immune response against Streptococcus,  
CC and to elicit a similar response to a target antigen fused to the  
CC protein. Unlike other immunological carriers, Hsp60 proteins are not  
CC immunosuppressive so provide an increased response to any conjugated or  
CC fused antigen. Also, where used for cancer control, they lack the side  
CC effects associated with endotoxins. They can also be used to detect  
CC specific antibodies and in treatment or prevention of tumours  
CC (e.g. sarcoma or cancers of breast, ovary, prostate, lung, pancreas or  
CC liver). The Hsp60 polynucleotide is used for recombinant production  
CC of the protein, as a source of primers and probes for detecting  
CC streptococci in standard hybridization/amplification assays, and  
CC therapeutically in gene therapy vectors.  
XX  
SQ Sequence 1661 BP; 512 A; 308 C; 407 G; 434 T; 0 other;

Query Match 63.6%; Score 14; DB 20; Length 1661;  
Best Local Similarity 63.6%; Pred. No. 1.3e+02;  
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 tgttgagannnnmttgata 22

DB 1163 TGTGGAGCTCTACTTGTGATA 1142  
|||||||  
RESULT 12  
AAC47974  
ID AAC47974 standard; DNA; 2660 BP.  
XX  
AC AAC47974;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 55799.  
XX  
KW Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway;  
KW metabolic pathway; promoter; termination sequence; ss.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135293.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.



PR 18-JUN-1999; 99US-0139455.  
 PR 18-JUN-1999; 99US-0139456.  
 PR 18-JUN-1999; 99US-0139457.  
 PR 18-JUN-1999; 99US-0139458.  
 PR 18-JUN-1999; 99US-0139459.  
 PR 18-JUN-1999; 99US-0139460.  
 PR 18-JUN-1999; 99US-0139461.  
 PR 18-JUN-1999; 99US-0139462.  
 PR 18-JUN-1999; 99US-0139463.  
 PR 18-JUN-1999; 99US-0139750.  
 PR 18-JUN-1999; 99US-0139763.  
 PR 21-JUN-1999; 99US-0139817.  
 PR 22-JUN-1999; 99US-0139899.  
 PR 23-JUN-1999; 99US-0140353.  
 PR 23-JUN-1999; 99US-0140354.  
 PR 24-JUN-1999; 99US-0140695.  
 PR 28-JUN-1999; 99US-0140823.  
 PR 29-JUN-1999; 99US-0140991.  
 PR 30-JUN-1999; 99US-0141287.  
 PR 01-JUL-1999; 99US-0141842.  
 PR 01-JUL-1999; 99US-0142154.  
 PR 02-JUL-1999; 99US-0142055.  
 PR 06-JUL-1999; 99US-0142390.  
 PR 08-JUL-1999; 99US-0142803.  
 PR 09-JUL-1999; 99US-0142920.  
 PR 12-JUL-1999; 99US-0142977.  
 PR 13-JUL-1999; 99US-0143542.  
 PR 14-JUL-1999; 99US-0143624.  
 PR 15-JUL-1999; 99US-0144005.  
 PR 16-JUL-1999; 99US-0144085.  
 PR 16-JUL-1999; 99US-0144086.  
 PR 19-JUL-1999; 99US-0144325.  
 PR 19-JUL-1999; 99US-0144331.  
 PR 19-JUL-1999; 99US-0144332.  
 PR 19-JUL-1999; 99US-0144333.  
 PR 19-JUL-1999; 99US-0144334.  
 PR 19-JUL-1999; 99US-0144335.  
 PR 20-JUL-1999; 99US-0144352.  
 PR 20-JUL-1999; 99US-0144632.  
 PR 20-JUL-1999; 99US-0144884.  
 PR 21-JUL-1999; 99US-0144884.  
 PR 21-JUL-1999; 99US-0145086.  
 PR 21-JUL-1999; 99US-0145088.  
 PR 22-JUL-1999; 99US-0145085.  
 PR 22-JUL-1999; 99US-0145087.  
 PR 22-JUL-1999; 99US-0145089.  
 PR 22-JUL-1999; 99US-0145192.  
 PR 23-JUL-1999; 99US-0145145.  
 PR 23-JUL-1999; 99US-0145218.  
 PR 23-JUL-1999; 99US-0145224.  
 PR 26-JUL-1999; 99US-0145276.  
 PR 27-JUL-1999; 99US-0145913.  
 PR 27-JUL-1999; 99US-0145918.  
 PR 27-JUL-1999; 99US-0145919.  
 PR 28-JUL-1999; 99US-0145931.  
 PR 02-AUG-1999; 99US-0146386.  
 PR 02-AUG-1999; 99US-0146388.  
 PR 03-AUG-1999; 99US-0146389.  
 PR 03-AUG-1999; 99US-0147038.  
 PR 04-AUG-1999; 99US-0147204.  
 PR 04-AUG-1999; 99US-0147302.  
 PR 05-AUG-1999; 99US-0147192.  
 PR 05-AUG-1999; 99US-0147260.  
 PR 06-AUG-1999; 99US-0147303.  
 PR 06-AUG-1999; 99US-0147416.  
 PR 09-AUG-1999; 99US-0147493.  
 PR 09-AUG-1999; 99US-0147933.  
 PR 10-AUG-1999; 99US-0148171.  
 PR 11-AUG-1999; 99US-0148319.  
 PR 12-AUG-1999; 99US-0148341.  
 PR 13-AUG-1999; 99US-0148565.  
 PR 13-AUG-1999; 99US-0148684.  
 PR 16-AUG-1999; 99US-0149368.

PR 17-AUG-1999; 99US-0149175.  
 PR 18-AUG-1999; 99US-0149426.  
 PR 20-AUG-1999; 99US-0149722.  
 PR 20-AUG-1999; 99US-0149723.  
 PR 20-AUG-1999; 99US-0149929.  
 PR 23-AUG-1999; 99US-0149902.  
 PR 23-AUG-1999; 99US-0149930.  
 PR 25-AUG-1999; 99US-0150566.  
 PR 26-AUG-1999; 99US-0150884.  
 PR 27-AUG-1999; 99US-0151065.  
 PR 27-AUG-1999; 99US-0151066.  
 PR 27-AUG-1999; 99US-0151080.  
 PR 30-AUG-1999; 99US-0151303.  
 PR 31-AUG-1999; 99US-0151438.  
 PR 01-SEP-1999; 99US-0151930.  
 PR 07-SEP-1999; 99US-0152363.  
 PR 10-SEP-1999; 99US-0153070.  
 PR 13-SEP-1999; 99US-0153758.  
 PR 15-SEP-1999; 99US-0154018.  
 PR 16-SEP-1999; 99US-0154039.  
 PR 20-SEP-1999; 99US-0154779.  
 PR 22-SEP-1999; 99US-0155139.  
 PR 23-SEP-1999; 99US-0155486.  
 PR 24-SEP-1999; 99US-0155659.  
 PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158023.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 13-OCT-1999; 99US-0159295.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161922.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

Query Match 63.6%; Score 14; DB 21; Length 2660;  
 Best Local Similarity 63.6%; Pred. NO. 1.4e+02;  
 Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 tggtagannnnnnnttgata 22  
 |||||  
 Db 955 tggtagagtcgtgggttgata 976

RESULT 13  
 AAH56861/C  
 ID AAH56861 standard; DNA: 2663 BP.

XX AAH56861;  
 AC  
 XX  
 DT 06-SEP-2001 (first entry)  
 XX  
 DE S. pyogenes groEL gene partial sequence SEQ ID NO:7.  
 XX  
 XX Antisense oligonucleotide; groE; groEL; groES; inhibitor; growth;  
 KM microorganism; Escherichia coli; Streptococcus pneumoniae; diagnosis;  
 KM Streptococcus pyogenes; Staphylococcus aureus; Pseudomonas aeruginosa;  
 KM antibacterial; antiviral; antiproliferative; antisense therapy;  
 KM microbial infection; ds.  
 XX  
 XX Streptococcus pyogenes.  
 OS  
 XX  
 PN MO200136625-A2.  
 PD  
 XX 25-MAY-2001.  
 XX  
 PF 20-NOV-2000; 2000MO-CA01347.  
 XX  
 PR 18-NOV-1999; 99US-0166249.  
 XX  
 PA (GENE-) GENESENSE TECHNOLOGIES INC.  
 XX  
 PI Wright JA, Young AH, Dugourd D;  
 XX  
 XX WPI; 2001-355633/37.  
 DR  
 XX  
 PT Novel antisense compounds targeting nucleic acid encoding groEL or  
 PT groES gene of microorganism, which hybridize with and inhibit  
 PT expression of the genes, useful to inhibit growth of microorganism  
 PT having the genes -  
 XX  
 PS Disclosure; Fig 7; 110pp; English.

XX The present invention specifically claims AAH56368 to AAH56832 which are  
 CC antisense oligonucleotides to nucleotide sequences encoding groE. More  
 CC generally, antisense compounds (I) comprising antisense oligonucleotides  
 CC of 5-50 bases targeted to a nucleotide sequence encoding groEL (heat  
 CC shock protein (HSP)60) (GL) and groES (HSP10) (GS) gene from a  
 CC microorganism, where the antisense compound is complementary to GL or  
 CC GS of a microorganism and specifically hybridizes with and inhibits the  
 CC expression of GL or GS, is claimed. (I) have antibacterial, antiviral  
 CC and antiproliferative activities, and can be used in antisense therapy  
 CC and for inhibition of expression of groES or groEL. (I) are useful for  
 CC inhibiting expression of GL or GS in cells or tissues in vitro. (I) are  
 CC also useful for inhibiting the growth of a microorganism, or inhibiting  
 CC the expression of GL or GS gene in a microorganism (a bacterial cell or  
 CC a virus) having a GL or GS gene which involves administering to the  
 CC microorganism or to a cell infected with the microorganism, (I). (I) are  
 CC also useful for treating a mammalian pathological condition mediated by  
 CC the microorganisms which involves identifying a eukaryotic organism  
 CC having a pathological condition mediated by microorganisms having a GL  
 CC or GS gene and administering (I) such that the growth of microorganism  
 CC is inhibited. The antisense compounds are utilized for diagnostics,  
 CC therapeutics, prophylaxis and as research reagents and kits, e.g., to  
 CC prevent or delay microbial infections in humans. They are also useful as  
 CC molecular weight markers. AAH56362 to AAH56367 and AAH56833 to AAH56854  
 CC represent PCR primers for groE sequences which are used in the  
 CC exemplification of the present invention. AAH56855 to AAH56870 represent  
 CC groE nucleotide sequence given in the present invention.  
 CC  
 XX  
 SQ Sequence 2683 BP; 808 A; 490 C; 593 G; 792 T; 0 other;

Query Match 63.6%; Score 14; DB 22; Length 2683;  
 Best Local Similarity 63.6%; Pred. No. 1.4e+02;  
 Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
 QY 1 tgttgannnnnnnttgata 22  
 DB 963 TGTGGAGCTCTACTTTGATA 942

RESULT 14  
 AA099430/C  
 ID AA099430 standard; DNA; 4197 BP.  
 XX  
 AC AA099430;  
 XX  
 DT 22-DEC-1995 (first entry)  
 XX  
 DE B. sphaericus SLP gene.  
 XX  
 KM Surface layer protein; SLP; fusion protein; vaccine; antigen;  
 KM surface expression; epitope; ds.  
 XX  
 OS Bacillus sphaericus.  
 XX  
 FH Key Location/Qualifiers  
 FT RBS 79..85  
 FT /\*tag= a  
 FT CDS 95..3853  
 FT /\*tag= b  
 FT sig-peptide 95..164  
 FT /\*tag= c  
 FT mat-peptide 185..3850  
 FT /\*tag= d  
 XX  
 PN MO9519371-A2.  
 PD  
 XX 20-JUL-1995.  
 XX  
 PF 13-JAN-1995; 95MO-EP00147.  
 XX  
 PR 14-JAN-1994; 94GB-0000650.  
 XX  
 PA (SOLV ) SOLVAY SA.  
 XX  
 PI Deblaere RY, Desomer J, Dhaese P;  
 XX  
 DR WPI; 1995-263827/34.  
 DR P-PADB; AAR80530.  
 XX  
 PT Host cell expressing surface layer protein fusion protein - used for  
 PT presentation of antigens and vaccine prodn.  
 XX  
 PS Disclosure; Fig.6; 95pp; English.

XX A probe based on the N-terminal sequence of B. sphaericus P-1 (LMG  
 CC P-13855) surface layer protein was used to screen an HindIII-  
 CC generated library to isolate the slp gene. Promoter regions  
 CC of the gene are used in genetic constructs providing surface  
 CC expression of heterologous proteins in P-1 hosts.  
 CC  
 XX  
 SQ Sequence 4197 BP; 1470 A; 731 C; 763 G; 1233 T; 0 other;

Query Match 63.6%; Score 14; DB 16; Length 4197;  
 Best Local Similarity 63.6%; Pred. No. 1.5e+02;  
 Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
 QY 1 tgttgannnnnnnttgata 22  
 DB 2836 TGTGGAGCAACTGTTTGATA 2815

RESULT 15  
 AAC28622  
 ID AAC28622 standard; cDNA; 149 BP.  
 XX  
 AC AAC28622;  
 XX  
 DT 06-OCT-2000 (first entry)  
 XX

```

DE Human secreted protein 5' EST, SEQ ID NO: 32697.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX
PN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
PR 26-FEB-1999; 99US-0122487.
XX
PA (GEST ) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI; 2000-500381/45.
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
PS
PS Claim 1: SEQ ID 32697; 71pp + CD-ROM; English.
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.
XX
SQ Sequence 149 BP; 33 A; 14 C; 37 G; 65 T; 0 other;

```

```

Query Match          59.1%; Score 13; DB 21; Length 149;
Best Local Similarity 61.9%; Pred. No. 3.4e+02;
Matches 13; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

```

Oy 1 tgttgannnnnnnlttgat 21
   |||||
DB 13 tgttgattctatttgat 33

```

Search completed: December 15, 2001, 03:29:46  
Job time: 7886 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 03:22:47 : Search time 172.39 Seconds  
(Without alignments)  
28.903 Million cell updates/sec

Title: US-09-380-826a-6

Perfect score: 1 lgltgannnnnnnttgata 22

Scoring table: IDENTITY-NUC

Gapop 10.0, Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_MN:\*

- 1: /cgn2\_6/ptodata/2/1na/5A-COMB.seq:\*
- 2: /cgn2\_6/ptodata/2/1na/5B-COMB.seq:\*
- 3: /cgn2\_6/ptodata/2/1na/6A-COMB.seq:\*
- 4: /cgn2\_6/ptodata/2/1na/6B-COMB.seq:\*
- 5: /cgn2\_6/ptodata/2/1na/PCBUS-COMB.seq:\*
- 6: /cgn2\_6/ptodata/2/1na/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysts of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
C 1	14	63.6	3666	2	US-08-682-517-13
C 2	14	63.6	3666	2	US-08-682-517-14
C 3	14	63.6	4197	2	US-08-682-517-7
C 4	14	63.6	4197	2	US-08-682-517-8
C 5	13	59.1	3487	4	US-09-276-531-32
C 6	13	59.1	5253	3	US-08-714-918-19
C 7	13	59.1	5253	4	US-09-265-315-19
C 8	13	59.1	5253	4	US-09-265-315-19
C 9	13	59.1	5253	4	US-09-266-417-19
C 10	12.4	56.4	1390	4	US-08-821-994-61
C 11	12.4	56.4	1434	4	US-08-821-994-62
C 12	12.4	56.4	1760	1	US-08-413-118-118
C 13	12.4	56.4	1760	3	US-08-473-446-118
C 14	12.4	56.4	1800	1	US-08-752-238-2
C 15	12.4	56.4	1800	3	US-09-085-603B-2
C 16	12.4	56.4	1800	3	US-09-031-897-6
C 17	12.4	56.4	2730	1	US-08-339-129-1
C 18	12.4	56.4	2761	1	US-08-752-238-1
C 19	12.4	56.4	2761	3	US-09-085-603B-1
C 20	12.4	56.4	2761	3	US-09-031-897-5
C 21	12.4	56.4	2761	4	US-09-257-770-5
C 22	12.4	56.4	3438	3	US-08-613-009A-1
C 23	12.4	56.4	3997	4	US-08-821-994-72
C 24	12.4	54.5	527	4	US-08-954-395A-19
C 25	12.4	54.5	1130	1	US-07-864-004B-1
C 26	12.4	54.5	1130	1	US-08-251-937A-1
C 27	12.4	54.5	1130	1	US-08-212-133A-5

28	12	54.5	1130	1	US-08-474-503-3	Sequence 3, Appl
29	12	54.5	1130	2	US-08-670-707A-3	Sequence 3, Appl
30	12	54.5	1130	4	US-09-037-601-3	Sequence 3, Appl
31	12	54.5	1130	5	PCT-US93-03275-1	Sequence 1, Appl
32	12	54.5	1130	5	PCT-US94-13200-3	Sequence 3, Appl
33	12	54.5	1202	4	US-09-293-322C-9	Sequence 9, Appl
34	12	54.5	1623	1	US-08-121-202-3	Sequence 3, Appl
35	12	54.5	2026	3	US-08-755-587-26	Sequence 26, Appl
C 36	12	54.5	2625	1	US-08-468-036-2	Sequence 2, Appl
C 37	12	54.5	2625	2	US-08-376-843-2	Sequence 2, Appl
C 38	12	54.5	3284	3	US-09-136-652-1	Sequence 1, Appl
C 39	12	54.5	4173	4	US-08-981-729-9	Sequence 9, Appl
40	12	54.5	4173	4	US-08-981-446B-2	Sequence 2, Appl
41	12	54.5	4334	2	US-08-670-707A-38	Sequence 38, Appl
42	12	54.5	4334	4	US-09-037-601-38	Sequence 38, Appl
C 43	12	54.5	4931	3	US-09-058-489-20	Sequence 20, Appl
44	12	54.5	6057	3	US-08-362-525-1	Sequence 1, Appl
45	12	54.5	6402	2	US-08-670-707A-36	Sequence 36, Appl

## ALIGNMENTS

```
RESULT 1
US-08-682-517-13/C
: Sequence 13, Application US/08682517
: Patent No. 5874267
: GENERAL INFORMATION:
: APPLICANT:
: TITLE OF INVENTION: Expression of surface layer proteins
: NUMBER OF SEQUENCES: 25
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/682,517
: FILING DATE:
: CLASSIFICATION:
: INFORMATION FOR SEQ ID NO: 13:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3666 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: US-08-682-517-13

Query Match 63.6%; Score 14; DB 2; Length 3666;
Best Local Similarity 63.6%; Pred. No. 32;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 lgltgannnnnnnttgata 22
Db 2652 TGTTCGACCACTGTTGATA 2631

RESULT 2
US-08-682-517-14/C
: Sequence 14, Application US/08682517
: Patent No. 5874267
: GENERAL INFORMATION:
: APPLICANT:
: TITLE OF INVENTION: Expression of surface layer proteins
: NUMBER OF SEQUENCES: 25
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
: CURRENT APPLICATION DATA:
```

```

; APPLICATION NUMBER: US/08/682,517
; FILING DATE:
; CLASSIFICATION:
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3666 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3666
; US-08-682-517-14

Query Match      63.6%; Score 14; DB 2; Length 3666;
Best Local Similarity 63.6%; Pred. No. 32;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 ttttgannnnnnnttgata 22
    ||||||| |||||||
Db 2652 tgttgagcactggttgata 2631

RESULT 3
US-08-682-517-7/C
; Sequence 7, Application US/08682517
; Patent No. 5874267
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Expression of surface layer proteins
; NUMBER OF SEQUENCES: 25
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/682,517
; FILING DATE:
; CLASSIFICATION:
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4197 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-682-517-7

Query Match      63.6%; Score 14; DB 2; Length 4197;
Best Local Similarity 63.6%; Pred. No. 33;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 ttttgannnnnnnttgata 22
    ||||||| |||||||
Db 2836 tgttgagcactggttgata 2815

RESULT 4
US-08-682-517-8/C
; Sequence 8, Application US/08682517
; Patent No. 5874267
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Expression of surface layer proteins
; NUMBER OF SEQUENCES: 25
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```

; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/682,517
; FILING DATE:
; CLASSIFICATION:
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4197 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Bacillus sphaericus
; INDIVIDUAL ISOLATE: P-1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 95..3850
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 185..3850
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 95..184
; US-08-682-517-8

Query Match      63.6%; Score 14; DB 2; Length 4197;
Best Local Similarity 63.6%; Pred. No. 33;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 ttttgannnnnnnttgata 22
    ||||||| |||||||
Db 2836 tgttgagcactggttgata 2815

RESULT 5
US-09-276-531-32/C
; Sequence 32, Application US/09276531
; Patent No. 6183968
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lai, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Reddy, Roopa
; APPLICANT: Guegler, Karl J.
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF GENES ENCODING
; RECEPTORS AND PROTEINS ASSOCIATED WITH CELL PROLIFERATION
; NUMBER OF SEQUENCES: 134
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/276,531
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/079,677
; FILING DATE: March 27, 1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lynn E. Murry, Ph.D.
```

```

: REGISTRATION NUMBER: 42,918
: REFERENCE/DOCKET NUMBER: PA-0008 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (650) 855-0555
: TELEFAX: (650) 845-4166
: INFORMATION FOR SEQ ID NO: 32:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3487 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: PROSTUT01
: CLONE: 714029
: US-09-276-531-32

```

```

Query Match 59.1%; Score 13; DB 4; Length 3487;
Best Local Similarity 61.9%; Pred. No. 1.1e+02;
Matches 13; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

```

Oy 1 ttttgannnnnnnttgat 21
Db 1880 GTTGGAGTCACTGTTGAT 1860

```

```

RESULT 6
US-08-714-918-19
: Sequence 19, Application US/08714918
: Patent No. 6037123
: GENERAL INFORMATION:
: APPLICANT: Benton, Bret
: APPLICANT: Lee, Ving
: APPLICANT: Malouin, Francois
: APPLICANT: Martin, Patrick K.
: APPLICANT: Schmid, Molly B.
: APPLICANT: Sun, Dongxu
: TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS ANTIBACTERIAL
: TITLE OF INVENTION: TARGET GENES
: NUMBER OF SEQUENCES: 111
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Lyon & Lyon
: STREET: 633 West Fifth Street
: STREET: Suite 4700
: CITY: Los Angeles
: STATE: California
: COUNTRY: U.S.A.
: ZIP: 90071-2066
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
: MEDIUM TYPE: storage
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: IBM P.C. DOS 5.0
: SOFTWARE: Word Perfect 5.1
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/714,918
: FILING DATE: September 13, 1996
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 60/009,102
: FILING DATE: December 22, 1995
: APPLICATION NUMBER: 60/003,798
: FILING DATE: September 15, 1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Warburg, Richard J.
: REGISTRATION NUMBER: 32,327
: REFERENCE/DOCKET NUMBER: 222/005
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (213) 489-1600
: TELEFAX: (213) 955-0440
: TELEX: 67-3510
: INFORMATION FOR SEQ ID NO: 19:
: SEQUENCE CHARACTERISTICS:

```

```

: LENGTH: 5253 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: US-08-714-918-19

```

```

Query Match 59.1%; Score 13; DB 3; Length 5253;
Best Local Similarity 61.9%; Pred. No. 1.2e+02;
Matches 13; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

```

Oy 2 gttgannnnnnnttgata 22
Db 2895 GTTGGATGATGCTTTGATA 2915

```

```

RESULT 7
US-09-265-315-19
: Sequence 19, Application US/09265315
: Patent No. 6187541
: GENERAL INFORMATION:
: APPLICANT: Benton, Bret
: APPLICANT: Lee, Ving J.
: APPLICANT: Malouin, Francois
: APPLICANT: Martin, Patrick K.
: APPLICANT: Schmid, Molly B.
: APPLICANT: Sun, Dongxu
: TITLE OF INVENTION: METHODS OF SCREENING FOR COMPOUNDS
: TITLE OF INVENTION: ACTIVE ON STAPHYLOCOCCUS AUREUS
: NUMBER OF SEQUENCES: 111
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Lyon & Lyon
: STREET: 633 West Fifth Street
: STREET: Suite 4700
: CITY: Los Angeles
: STATE: California
: COUNTRY: U.S.A.
: ZIP: 90071-2066
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
: MEDIUM TYPE: storage
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: IBM P.C. DOS 5.0
: SOFTWARE: Word Perfect 5.1
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/265,315
: FILING DATE: March 9, 1999
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/714,918
: FILING DATE: September 13, 1996
: APPLICATION NUMBER: 60/009,102
: FILING DATE: December 22, 1995
: APPLICATION NUMBER: 60/003,798
: FILING DATE: September 15, 1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Warburg, Richard J.
: REGISTRATION NUMBER: 32,327
: REFERENCE/DOCKET NUMBER: 240/247
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (213) 489-1600
: TELEFAX: (213) 955-0440
: TELEX: 67-3510
: INFORMATION FOR SEQ ID NO: 19:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 5253 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: US-09-265-315-19

```

Query Match 59.1%; Score 13; DB 4; Length 5253;  
Best Local Similarity 61.9%; Pred. No. 1.2e+02;  
Matches 13; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 2 gtggannnnnnnttgata 22  
||||| |||||||  
Db 2895 GTGGATGTAGTCTTTTGATA 2915

## RESULT 8

US-09-265-315-19  
; Sequence 19, Application US/09265315  
; Patent No. 6187541  
; GENERAL INFORMATION:  
; APPLICANT: Benton, Bret  
; APPLICANT: Lee, Ving J.  
; APPLICANT: Malouin, Francois  
; APPLICANT: Martin, Patrick K.  
; APPLICANT: Schmid, Molly B.  
; APPLICANT: Sun, Dongxu  
; TITLE OF INVENTION: METHODS OF SCREENING FOR COMPOUNDS  
; TITLE OF INVENTION: ACTIVE ON STAPHYLOCOCCUS AUREUS  
; TITLE OF INVENTION: TARGET GENES  
; NUMBER OF SEQUENCES: 111  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; STREET: Suite 4700  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90071-2066  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; MEDIUM TYPE: Storage  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: IBM P.C. DOS 5.0  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/265,315  
; FILING DATE: March 9, 1999  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/714,918  
; FILING DATE: September 13, 1996  
; APPLICATION NUMBER: 60/009,102  
; FILING DATE: December 22, 1995  
; APPLICATION NUMBER: 60/003,798  
; FILING DATE: September 15, 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warburg, Richard J.  
; REGISTRATION NUMBER: 32,327  
; REFERENCE/DOCKET NUMBER: 240/247  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (213) 489-1600  
; TELEFAX: (213) 955-0440  
; TELEX: 67-3510  
; INFORMATION FOR SEQ ID NO: 19:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5253 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-09-265-315-19

Query Match 59.1%; Score 13; DB 4; Length 5253;  
Best Local Similarity 61.9%; Pred. No. 1.2e+02;  
Matches 13; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 2 gtggannnnnnnttgata 22  
||||| |||||||  
Db 2895 GTGGATGTAGTCTTTTGATA 2915

RESULT 9  
US-09-266-417-19

; Sequence 19, Application US/09266417  
; Patent No. 6228588  
; GENERAL INFORMATION:  
; APPLICANT: Benton, Bret  
; APPLICANT: Lee, Ving J.  
; APPLICANT: Malouin, Francois  
; APPLICANT: Martin, Patrick K.  
; APPLICANT: Schmid, Molly B.  
; APPLICANT: Sun, Dongxu  
; TITLE OF INVENTION: METHODS OF SCREENING FOR COMPOUNDS  
; TITLE OF INVENTION: ACTIVE ON STAPHYLOCOCCUS AUREUS  
; TITLE OF INVENTION: TARGET GENES  
; NUMBER OF SEQUENCES: 111  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; STREET: Suite 4700  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 90071-2066  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; MEDIUM TYPE: Storage  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: IBM P.C. DOS 5.0  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/266,417  
; FILING DATE: March 9, 1999  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/714,918  
; FILING DATE: September 13, 1996  
; APPLICATION NUMBER: 60/009,102  
; FILING DATE: December 22, 1995  
; APPLICATION NUMBER: 60/003,798  
; FILING DATE: September 15, 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warburg, Richard J.  
; REGISTRATION NUMBER: 32,327  
; REFERENCE/DOCKET NUMBER: 240/248  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (213) 489-1600  
; TELEFAX: (213) 955-0440  
; TELEX: 67-3510  
; INFORMATION FOR SEQ ID NO: 19:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5253 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-09-266-417-19

Query Match 59.1%; Score 13; DB 4; Length 5253;  
Best Local Similarity 61.9%; Pred. No. 1.2e+02;  
Matches 13; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 2 gtggannnnnnnttgata 22  
||||| |||||||  
Db 2895 GTGGATGTAGTCTTTTGATA 2915

RESULT 10  
US-08-821-994-61/c  
; Sequence 61, Application US/08821994A  
; Patent No. 6228643  
; GENERAL INFORMATION:



```

: APPLICANT: Greenland, Andrew J
: APPLICANT: Thomas, Didier RP
: TITLE OF INVENTION: Promoters
: FILE REFERENCE: PPD 50108
: CURRENT APPLICATION NUMBER: US/08/821,994A
: EARLIER FILING DATE: 1997-03-22
: EARLIER APPLICATION NUMBER: PCT/GB97/00729
: EARLIER FILING DATE: 1997-03-18
: EARLIER APPLICATION NUMBER: GB 9606062.9
: NUMBER OF SEQ ID NOS: 89
: SOFTWARE: Patentln Ver. 2.1
: SEQ ID NO 61
: LENGTH: 1390
: TYPE: DNA
: ORGANISM: Brassica napus
US-08-821-994-61

```

```

Query Match          56.4%; Score 12.4; DB 4; Length 1390;
Best Local Similarity 59.1%; Pred. No. 2.1e+02;
Matches 13; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```

```

OY 1 tgttgannnnnnnnttgata 22
    |||||
Db 26 TGTGCTTTCTAGTTTGATA 5

```

```

RESULT 11
US-08-821-994-62/C
: Sequence 62, Application US/08821994A
: Patent No. 6228643
: GENERAL INFORMATION:
: APPLICANT: Greenland, Andrew J
: APPLICANT: Thomas, Didier RP
: APPLICANT: Jepson, Ian
: TITLE OF INVENTION: Promoters
: FILE REFERENCE: PPD 50108
: CURRENT APPLICATION NUMBER: US/08/821,994A
: EARLIER FILING DATE: 1997-03-22
: EARLIER APPLICATION NUMBER: PCT/GB97/00729
: EARLIER FILING DATE: 1997-03-18
: EARLIER APPLICATION NUMBER: GB 9606062.9
: EARLIER FILING DATE: 1996-03-22
: NUMBER OF SEQ ID NOS: 89
: SOFTWARE: Patentln Ver. 2.1
: SEQ ID NO 62
: LENGTH: 1434
: TYPE: DNA
: ORGANISM: Brassica napus
US-08-821-994-62

```

```

Query Match          56.4%; Score 12.4; DB 4; Length 1434;
Best Local Similarity 59.1%; Pred. No. 2.1e+02;
Matches 13; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```

```

OY 1 tgttgannnnnnnnttgata 22
    |||||
Db 26 TGTGCTTTCTAGTTTGATA 5

```

```

RESULT 12
US-08-413-118-118
: Sequence 118, Application US/08413118
: Patent No. 5688920
: GENERAL INFORMATION:
: APPLICANT: PAOLETTI, ENZO
: APPLICANT: LIMBACH, KEITH J.
: TITLE OF INVENTION: NOCLOETIDE AND AMINO ACID SEQUENCES OF
: CANINE HERPESVIRUS 9B, 9C, AND 9D AND USES THEREFOR
: NUMBER OF SEQUENCES: 128

```

```

CORRESPONDENCE ADDRESS:
ADDRESS: CURTIS, MORRIS & SAFORD, P.C.
STREET: 530 FIFTH AVENUE, 25TH FLOOR
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: UNITED STATES OF AMERICA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/413,118
FILING DATE: 29-MAR-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/220,151
FILING DATE: 30-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: FROMMER, WILLIAM S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2670
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 118:
SEQUENCE CHARACTERISTICS:
LENGTH: 1760 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-413-118-118

```

```

Query Match          56.4%; Score 12.4; DB 1; Length 1760;
Best Local Similarity 59.1%; Pred. No. 2.2e+02;
Matches 13; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```

```

OY 1 tgttgannnnnnnnttgata 22
    |||||
Db 1076 TGTGATAACTATTTTATTA 1097

```

```

RESULT 13
US-08-473-446-118
: Sequence 118, Application US/08473446
: Patent No. 6017542
: GENERAL INFORMATION:
: APPLICANT: PAOLETTI, ENZO
: APPLICANT: LIMBACH, KEITH J.
: TITLE OF INVENTION: NOCLOETIDE AND AMINO ACID SEQUENCES OF
: CANINE HERPESVIRUS 9B, 9C, AND 9D AND USES THEREFOR
: NUMBER OF SEQUENCES: 128
CORRESPONDENCE ADDRESS:
ADDRESS: CURTIS, MORRIS & SAFORD, P.C.
STREET: 530 FIFTH AVENUE, 25TH FLOOR
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: UNITED STATES OF AMERICA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/473,446
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/413,118

```

```

: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: FROMMER, WILLIAM S.
: REGISTRATION NUMBER: 25,506
: REFERENCE/DOCKET NUMBER: 454310-2670
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 840-3333
: TELEFAX: (212) 840-0712
: INFORMATION FOR SEQ ID NO: 118:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1760 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: US-08-473-446-118

Query Match      56.4%; Score 12.4; DB 3; Length 1760;
Best Local Similarity 59.1%; Pred. No. 2.2e+02;
Matches 13; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY      1 ttttgagannnnnnnttgata 22
Db      1076 TGTGGATTAATAATTTTGATA 1097

RESULT 14
US-08-752-238-2/C
: Sequence 2, Application US/08752238
: Patent No. 5804418
: GENERAL INFORMATION:
: APPLICANT: Lambowitz Dr., Alan M
: APPLICANT: Mohr Dr., Georg
: APPLICANT: Saldanha Dr., Roland
: APPLICANT: Matsura Dr., Manabu
: TITLE OF INVENTION: Method for Preparing Nucleotide
: NUMBER OF INVENTION: 12
: NUMBER OF SEQUENCES: 12
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: CALFEE, HALTER & GRISWOLD
: STREET: 800 Superior Avenue
: CITY: Cleveland
: STATE: Ohio
: COUNTRY: US
: ZIP: 44114
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/752,238
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Goitick, Mary E.
: REGISTRATION NUMBER: 34,829
: REFERENCE/DOCKET NUMBER: 24671/00103
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (216) 622-8458
: TELEFAX: (216) 241-0816
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1800 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..1800
: US-08-752-238-2

```

```

Query Match      56.4%; Score 12.4; DB 1; Length 1800;
Best Local Similarity 59.1%; Pred. No. 2.2e+02;
Matches 13; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY      1 ttttgagannnnnnnttgata 22
Db      139 TATTGGAATATAAATTTTGATA 118

RESULT 15
US-09-085-603B-2/C
: Sequence 2, Application US/09085603B
: Patent No. 6001608
: GENERAL INFORMATION:
: APPLICANT: Lambowitz Dr., Alan M
: APPLICANT: Mohr Dr., Georg
: APPLICANT: Saldanha Dr., Roland
: APPLICANT: Matsura Dr., Manabu
: APPLICANT: Yang Dr., Jiam
: APPLICANT: Zimmerly Dr., Steven
: APPLICANT: Guo Dr., Huatao
: APPLICANT: Beall Dr., Clifford J.
: TITLE OF INVENTION: Methods of Making an Rnp Particle
: NUMBER OF INVENTION: Having Nucleotide Integrase Activity
: NUMBER OF SEQUENCES: 23
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: CALFEE, HALTER & GRISWOLD LLP
: STREET: 800 Superior Avenue
: CITY: Cleveland
: STATE: Ohio
: COUNTRY: US
: ZIP: 44114
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/085,603B
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Docherly, Pamela A.
: REGISTRATION NUMBER: 40,591
: REFERENCE/DOCKET NUMBER: 24671/04000
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (216) 241-0816
: TELEFAX: (216) 622-8416
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1800 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..1800
: US-09-085-603B-2

Query Match      56.4%; Score 12.4; DB 3; Length 1800;
Best Local Similarity 59.1%; Pred. No. 2.2e+02;
Matches 13; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY      1 ttttgagannnnnnnttgata 22
Db      139 TATTGGAATATAAATTTTGATA 118

```

Search completed: December 15, 2001, 03:22:49  
Job time: 7564 sec

---

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 02:33:44 ; Search time 4441.54 Seconds  
(without alignments)  
53.226 Million cell updates/sec

Title: US-09-380-826A-6

Perfect score: 22

Sequence: 1 ttttgagannnnnnnttgata 22

Scoring table:

IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 11351937 segs, 537289281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*  
1: em\_estfun:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estom:\*  
5: em\_estopl:\*  
6: em\_estlda:\*  
7: em\_estro:\*  
8: em\_estov:\*  
9: em\_hic:\*  
10: qb\_estl1:\*  
11: qb\_estl2:\*  
12: qb\_hic:\*  
13: qb\_gss:\*  
14: em\_gss\_fun:\*  
15: em\_gss\_hum:\*  
16: em\_gss\_ino:\*  
17: em\_gss\_pln:\*  
18: em\_gss\_pro:\*  
19: em\_gss\_rod:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
C 1	14	63.6	184	10	BB357781 BB357781
C 2	14	63.6	210	10	BE059640 sn35a01.y
C 3	14	63.6	228	11	BE956026 UI-M-BH4-
C 4	14	63.6	235	10	BB170254 BB170254
C 5	14	63.6	235	13	AZ925147
C 6	14	63.6	250	10	A1122094 uc46f10.r
C 7	14	63.6	261	10	AM879900 QV3-OT002
C 8	14	63.6	265	10	BB287901 BB287901
C 9	14	63.6	266	11	BR800082 MR1-CT002
C 10	14	63.6	293	11	BE634128 NF082F04D
C 11	14	63.6	298	10	AA515145 ng68c01.s
C 12	14	63.6	332	10	BE202756 EST402778

C 13	14	63.6	336	13	AZ037213
C 14	14	63.6	342	13	BH083614
C 15	14	63.6	349	10	AU112918
C 16	14	63.6	359	10	AV546877
C 17	14	63.6	364	10	AM226052
C 18	14	63.6	371	13	AQ311167
C 19	14	63.6	376	13	AQ844476
C 20	14	63.6	376	13	AQ102282
C 21	14	63.6	378	10	AM240466
C 22	14	63.6	386	11	BF318641
C 23	14	63.6	389	10	AA488558
C 24	14	63.6	399	11	R89220
C 25	14	63.6	410	10	AM624588
C 26	14	63.6	410	13	AQ150667
C 27	14	63.6	416	10	AA555292
C 28	14	63.6	416	13	AQ345780
C 29	14	63.6	427	10	A1662442
C 30	14	63.6	433	11	BE464046
C 31	14	63.6	436	11	BE463084
C 32	14	63.6	447	13	AZ273408
C 33	14	63.6	459	11	BE854379
C 34	14	63.6	467	10	AL370583
C 35	14	63.6	471	10	A1591928
C 36	14	63.6	476	13	AQ914928
C 37	14	63.6	481	11	T75954
C 38	14	63.6	482	13	B29252
C 39	14	63.6	485	11	BE226581
C 40	14	63.6	489	10	BE196823
C 41	14	63.6	490	10	BE448206
C 42	14	63.6	493	10	AA637362
C 43	14	63.6	496	13	AQ589820
C 44	14	63.6	497	10	BE53412
C 45	14	63.6	498	11	BE003327

## ALIGNMENTS

RESULT 1 BB357781 184 bp mRNA EST 12-JUL-2000  
LOCUS BB357781 RIKEN full-length enriched, adult male corpus striatum Mus  
DEFINITION musculus cDNA clone C030024G02 3', mRNA sequence.  
ACCESSION BB357781  
VERSION BB357781.1 GI:9069609  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 184)  
AUTHORS Kono, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T., Carninci,  
P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N.,  
Hirozane, T., Hori, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M.,  
Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N.,  
Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C., Kusakabe, M.,  
Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y.,  
Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata,  
Y., Shigemoto, Y., Shinagawa, A., Shitaki, T., Sogabe, Y., Sugihara, Y.,  
Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomioka, N., Toya,  
T., Tsunoda, Y., Watanabe, A., Watanabe, S., Yamamura, T., Yamana, I.,  
Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino,  
M., Muramatsu, M. and Hayashizaki, Y.  
RIKEN Mouse ESTs (Kono, H., et al.)  
Unpublished (2000)  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center(GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22, Saitama-cho, Tsukuba, Ibaraki, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@gsc.riken.go.jp,

URL: <http://genome.gsc.riken.go.jp/>  
 Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S., Sasaki, N., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.  
 Thermostabilization and thermoactivation of thermostable enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)  
 Itoh, M., Katsunai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, Y., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y., and Hayashizaki, Y.  
 Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)  
 Carninci, P. and Hayashizaki, Y.  
 High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)  
 Please visit our web site (<http://genome.rtc.riken.go.jp/>) for further details.

## FEATURES

## source

Location/Qualifiers  
 1. 184  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="C030024G02"  
 /clone\_lib="RIKEN full-length enriched, adult male corpus striatum"  
 /sex="male"  
 /tissue\_type="corpus striatum"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /note="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGATCCAGACCTCTTTTCTTTTCTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 165.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATCTCCAGTTAATTAATTAATCCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified plasmid KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI"

BASE COUNT 51 a 34 c 39 g 60 t  
 ORIGIN

Query Match 63.6%; Score 14; DB 10; Length 184;  
 Best Local Similarity 63.6%; Pred. No. 2.1e+03;  
 Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Oy 1 tcttgagannnnnnnnttgata 22  
 |||||  
 Db 82 TGTGGAATGCTTTTGATA 61

RESULT 2  
 BE059640 210 bp mRNA EST 21-NOV-2000  
 LOCUS BE059640  
 DEFINITION sn35a01.v1 Gm-c1016 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-c1016-13441 5' similar to TR:092V13 092V13 PUTATIVE ELONGATION FACTOR.; mRNA sequence.

ACCESSION BE059640  
 VERSION BE059640.1 GI:8404006  
 KEYWORDS EST.  
 SOURCE soybean.  
 ORGANISM Glycine max  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;

## REFERENCE

## AUTHORS

TITLE  
 JOURNAL  
 COMMENT  
 Public Soybean EST Project  
 Unpublished (1999)  
 Contact: Shoemaker R/Public Soybean EST Project  
 Washington University School of Medicine  
 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: estelw@wustl.edu

Glycine.  
 1 (bases 1 to 210)  
 Shoemaker, R., Kaim, P., Vodkin, L., Erpelting, J., Corvelli, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.  
 Public Soybean EST Project  
 Unpublished (1999)  
 Contact: Shoemaker R/Public Soybean EST Project  
 Washington University School of Medicine  
 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: estelw@wustl.edu  
 This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134 For further information call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314) 427-3324 or contact: clones@genomesystems.com or info@genomesystems.com web site: www.genomesystems.com  
 Insert Length: 647 Std Error: 0.00  
 Seq primer: -40RP from Gibco  
 High quality sequence stop: 165.

## FEATURES

## source

Location/Qualifiers  
 1. 210  
 /organism="Glycine max"  
 /db\_xref="taxon:3847"  
 /clone="GENOME SYSTEMS CLONE ID: Gm-c1016-13441"  
 /clone\_lib="Gm-c1016"  
 /tissue\_type="Immature flowers of field grown plants"  
 /lab\_host="X110-Gold"  
 /note="Vector: plasmid II XRI; Site 1: EcoRI; Site 2: XhoI; This cDNA library was constructed from mRNA isolated from immature flowers of field grown plants. The cDNA library was prepared using the StrataGene plasmid II XR library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly (dT) sequence to the blunt-ended cDNA fragments followed by were ligated to the blunt-ended cDNA fragments directionally. XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the plasmid vector. The ligated cDNA fragments were transformed into X110-Gold host cells. This library was constructed by Dr. Randy Shoemaker and Dr. John Erpelting."

BASE COUNT 57 a 38 c 55 g 60 t  
 ORIGIN

Query Match 63.6%; Score 14; DB 10; Length 210;  
 Best Local Similarity 63.6%; Pred. No. 2.1e+03;  
 Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Oy 1 tcttgagannnnnnnnttgata 22  
 |||||  
 Db 43 TGTGATCTCAACGTTTGATA 64

RESULT 3  
 BE956026/c 228 bp mRNA EST 04-OCT-2000  
 LOCUS BE956026  
 DEFINITION UI-M-BH4-bav-a-11-0-UI.s1 NIH\_BMAP\_M.S5 Mus musculus cDNA clone UI-M-BH4-bav-a-11-0-UI 3', mRNA sequence.

ACCESSION BE956026  
 VERSION BE956026.1 GI:10600157  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 228)

**AUTHORS** Bonaldo,M.F., Lennon,G. and Soares,M.B.  
**TITLE** Normalization and subtraction: two approaches to facilitate gene discovery  
**JOURNAL** Genome Res. 6 (9), 791-806 (1996)  
**MEDLINE** 97044477  
**COMMENT** Contact: Chln, H  
 National Institute of Mental Health  
 6001 Executive Blvd., Room 7N-7190, MSC 9643, Bethesda, MD  
 20892-9643, USA  
 Tel: 301 443 1706  
 Fax: 301 443 9890  
 Email: mestr@mail.nih.gov

The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to identify it as a clone from the olfactory bulbs tissue cDNA Library Preparation: M.B. Soares Lab cDNA distribution: Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It should be noted that Bento Soares is generating a small number of additional specialized non-redundant arrays of BMAP cDNAs whose availability will be considered under appropriate and limited collaborative arrangements  
 Seq primer: M13 Forward  
 POLYA=Yes.

#### FEATURES

**source** 1. 228  
 Location/Qualifiers  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="UI-M-BH4-Dav-a-11-0-UI"  
 /clone.lib="NIH-BMAP\_M\_S5"  
 /dev\_stage="27-32 days"  
 /lab\_host="DH10B (Life Technologies)"  
 /note="Vector: pT7SD-Pac (Pharmacia) with a modified polylinker. Site\_1: Not I; Site\_2: Eco RI; The NIH-BMAP\_M\_S5 library is a subtracted library of a series, ultimately derived from a mixture of individually tagged normalized libraries from ten regions of the mouse brain (cerebellum, brain stems, olfactory bulbs, hypothalamus, cortex, amygdala, basal ganglia, pineal gland, striatum, hippocampus) after a series of subtractions to reduce the representation of cDNAs from which ESTs had already been generated. For a detailed description of the library from which this clone was derived, please visit our web site at [brainst.legu.wiowa.edu](http://brainst.legu.wiowa.edu).  
 TAG\_LIB=NIH-BMAP\_M\_S5  
 TAG\_RISUB="olfactory-bulbs"  
 TAG\_SEQ="CAAGG"

**BASE COUNT** 64 a 47 c 42 g 75 t  
**ORIGIN**

**Query Match** 63.6%; Score 14; DB 11; Length 228;  
**Best Local Similarity** 63.6%; Pred. NO. 2.1e+03;  
**Matches** 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

**QY** 1 tcttgannnnnnnttgata 22  
 |||||  
**Db** 144 tcttgaaatctgctttgata 123

**RESULT** 4  
**LOCUS** BB170254/c 235 bp mRNA EST 29-JUN-2000  
**DEFINITION** BB170254 RIKEN full-length enriched, adult male hypothalamus Mus  
**ACCESSION** BB170254  
**VERSION** BB170254.1 GI:8829337  
**KEYWORDS** EST.  
**SOURCE** house mouse.  
**ORGANISM** Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

#### REFERENCE

**AUTHORS** Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 235)  
 Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci,P., Endo,T., Fukuda,S., Fukunishi,Y., Haru,A., Hayatsu,N., Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N., Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M., Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y., Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y., Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tomimaga,N., Toya,T., Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yamanaka,I., Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.  
 RIKEN Mouse ESTs (Konno,H., et al.)  
 Unpublished (2000)

#### JOURNAL

**COMMENT** Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Saitoh-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@gsc.riken.go.jp,  
 URL:<http://genome.gsc.riken.go.jp/>  
 Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoaka,S., Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
 The mostablation and thermocatalysis of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)  
 Itoh,M., Kitsuana,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J., Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.  
 Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)  
 Carninci,P. and Hayashizaki,Y.  
 High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)  
 Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

#### FEATURES

**source** 1. 235  
 Location/Qualifiers  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="A230017D13"  
 /clone.lib="RIKEN full-length enriched, adult male hypothalamus"  
 /sex="male"  
 /tissue\_type="hypothalamus"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /note="Site\_1: SalI; Site\_2: BamHI. cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGATCCAGAGCTTTTGTTTTGTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATTCGACTTATTAATTAATCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pbluescript KS(+) after bulk excision from Lambda FLC I."

#### BASE COUNT

84 a 47 c 31 g 73 t

**Query Match**

63.6%; Score 14; DB 10; Length 235;

Best Local Similarity 63.6%; Pred. No. 2.1e+03;  
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 tcttggannnnnnnnttgata 22  
|||||||  
Db 66 TGTGGACATTGACTCTTTGATA 45

RESULT 5  
A2925147 235 bp DNA GSS 01-APR-2001  
LOCUS A2925147/c  
DEFINITION 4910.ez32e09.s1 Saccharomyces paradoxus N17 Saccharomyces paradoxus  
genomic clone 4910.ez32e09.s1, DNA sequence.

ACCESSION A2925147 GI:13496046  
VERSION A2925147.1  
KEYWORDS GSS.  
SOURCE Saccharomyces paradoxus.  
ORGANISM Saccharomyces paradoxus  
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
Saccharomycetales; Saccharomycetaceae; Saccharomyces.

REFERENCE 1 (bases 1 to 235)  
AUTHORS Clifton, P.F., Hillier, L.W., Fulton, L., Graves, T., Miner, T., Gish  
TITLE 'W.R., Waterston, R.H. and Johnston, M.  
JOURNAL Surveying Saccharomyces genomes to identify functional elements by  
COMMENT Comparative DNA sequence analysis  
Unpublished (2001)  
Contact: Johnston M  
Department of Genetics  
Washington University Medical School  
Box 8232, 4566 Scott Ave., St. Louis, MO 63110, USA  
Tel: 314 362 2735  
Fax: 314 362 7835  
Email: mjgenetics.wustl.edu  
Class: random plasmid subclone.

FEATURES  
Location/Qualifiers  
1..235  
/organism="Saccharomyces paradoxus"  
/strain="N17"  
/db\_xref="taxon:27291"  
/clone="4910.ez32e09.s1"  
/clone\_lib="Saccharomyces paradoxus N17"  
/note="Random genomic sequence"

BASE COUNT 79 a 35 c 36 g 85 t  
ORIGIN

Query Match 63.6%; Score 14; DB 13; Length 235;  
Best Local Similarity 63.6%; Pred. No. 2.1e+03;  
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 tcttggannnnnnnnttgata 22  
|||||||  
Db 232 TGTGGACATTGACTCTTTGATA 211

RESULT 6  
A1122094 250 bp mRNA EST 02-SEP-1998  
LOCUS A1122094  
DEFINITION uc46110.1 Soares\_mammary\_gland\_NMLMG Mus musculus CDNA clone  
IMAGE:1401067 5 Similar to SW:Y335\_MYCGE P47577 HYPOPHYETICAL  
GTP-BINDING PROTEIN MG335.1, mRNA sequence.

ACCESSION A1122094  
VERSION A1122094.1 GI:3522418  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 250)  
AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,  
Geisels, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,  
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,  
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and

TITLE Waterston, R.  
JOURNAL The WashU-HMI Mouse EST Project  
COMMENT Unpublished (1996)  
Contact: Maria M/Mouse EST Project  
WashU-HMI Mouse EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810

Email: mousestewartson.wustl.edu  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.lnl.gov) for further information.  
MGI:912783  
Seq primer: -28n13 rev2 ET from Amersham  
High quality sequence stop: 225.

FEATURES  
Location/Qualifiers  
1..250  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="IMAGE:1401067"  
/clone\_lib="Soares\_mammary\_gland\_NMLMG"  
/sex="female (lactating)"  
/tissue\_type="mammary gland"  
/lab\_host="DH10B"  
/note="Vector: p773P-Pac (Pharmacia) with a modified  
polylinker; 1st strand cDNA was prepared from mammary  
gland tissue from a lactating female, and was then primed  
with a Not I - oligo(dT) primer. Double-stranded cDNA was  
ligated to Eco RI adaptors (Pharmacia), digested with Not  
I and cloned into the Not I and Eco RI sites of the  
modified p773 vector. Library is normalized. Library  
was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 48 a 79 c 75 g 48 t  
ORIGIN

Query Match 63.6%; Score 14; DB 10; Length 250;  
Best Local Similarity 63.6%; Pred. No. 2.1e+03;  
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 tcttggannnnnnnnttgata 22  
|||||||  
Db 211 TGTGGAAATCCCTCTTTGATA 232

RESULT 7  
AM879900/c 261 bp mRNA EST 23-MAY-2000  
LOCUS AM879900  
DEFINITION OV3-OT0029-290300-134-b02 OT0029 Homo sapiens CDNA, mRNA sequence.  
ACCESSION AM879900  
VERSION AM879900.1 GI:8041910  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 261)  
Dias Neto, E., Garcia Correa, R., Verjowski-Almeida, S., Briones, M.R.,  
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,  
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare  
, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
Simpson, A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL MEDLINE  
COMMENT Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922



**FEATURES**  
**SOURCE**

BASE COUNT	79 a	84 c	30 g	68 f
ORIGIN				

```
QY      1  tgttgannnnnnnttgata 22
          |||||
Db      65  TGTTCGACTCAGAGTTTGATA 44
```

ACCESSION	BB287901	GI:8988350
VERSION	BB287901.1	
KEYWORDS	EST.	
SOURCE	house mouse.	
ORGANISM	Mus musculus	

TITLE  
JOURNAL  
COMMENT

P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N.,  
Hirozane, T., Horii, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M.,  
Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N.,  
Kiyosawa, H., Kojima, T., Kondo, S., Koya, S., Kurihara, C., Kusakabe, M.,  
Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y.,  
Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata, Y.,  
Shigemoto, Y., Shingawa, A., Shiraki, T., Sogabe, Y., Sugahara, Y.,  
Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomioka, N., Toyai,  
T., Tsunoda, Y., Watahiki, A., Watanabe, S., Yamamura, T., Yamanka, J.,  
Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino,  
M., Muramatsu, M. and Hayashizaki, Y.  
RIKEN Mouse ESTs (Kono, H., et al.).  
Unpublished (2000)  
Contact: Yoshinide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel.: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@gs.riken.go.jp,  
UR: <http://genome-gsc.riken.go.jp/>  
Carninci, F., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S., Sasakawa, N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Thermostabilization and thermooxidation of thermolabile enzymes by

**FEATURES**  
**SOURCE**

```

/note="Site.1: Sali; site.2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGGAGAGAGAAGATCCACAGAGCTCTTTTTTTTTTTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5'
GAGCGAGAGATTTCGCAGATTAAATTAATTAATCCCCCCCCCCC 3']. cDNA
was cleaved with XhoI and BamHI. Vector: a modified
pBluescript KS(+) after bulk excision from Lambda FLX I."

```

```
Qy      1  tgttgannnnnnnttgata  22
          |||||
Db      165 TGTGGAAATGGCTTTTGATA  144
```

RESULT	9
BFB00082/c	
LOCUS	BFB00082
DEFINITION	MRI-CI0021-161000-004-f02 CI0021 Homo sapiens EST
ACCESSION	BFB00082
VERSION	BFB00082.1
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens

REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	COMMENT
Eukaryota; Metazoa; Chordata; Crinata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Carnivora; Homiidae; Homo.	Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsumura, A., Bala, G.S., Simpson, D.H., Brito, A.A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags	Proc. Natl. Acad. Sci. U.S.A.	97 (7), 3491-3496 (2000)	200202653
	Contact: Simpson A.J.G.				
	Laboratory of Cancer Genetics				

Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=MR1&L2=MR1-C10021-  
161000-004-f02&t3=2000-10-16&t4=1)

Seq primer: puc 18 forward  
High quality sequence start: 12  
High quality sequence stop: 269.  
Location/Qualifiers

FEATURES  
source

1.269  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="C10021"  
/dev\_stage="Adult"  
/note="Organ: colon\_ins; Vector: puc18; Site:1: SmaI;  
Site:2: SmaI; A mini-library was made by cloning products  
derived from ORESTES PCR (U.S. Letters Patent application  
No. 196/716 - Ludwig Institute for Cancer Research)  
profiles into the puc 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."

BASE COUNT 88 a 53 c 44 g 84 t  
ORIGIN

Query Match 63.6%; Score 14; DB 11; Length 269;  
Best Local Similarity 63.6%; Pred. No. 2.1e+03;  
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Oy 1 tggtagannnnnnnttgata 22  
|||||||  
Db 93 tcttgcaatttcctcatttgata 72

RESULT 10  
BF634128 293 bp mRNA EST 19-DEC-2000  
LOCUS BF634128/C  
DEFINITION NP082F04D1F1042 Drought Medicago truncatula cDNA clone NF082F04DT  
5' mRNA sequence.  
ACCESSION BF634128  
VERSION BF634128.1 GI:11898286  
KEYWORDS EST.  
SOURCE barrel medic.  
ORGANISM Medicago truncatula  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;  
Medicago.  
1 (bases 1 to 293)  
Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,  
Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.  
Expressed Sequence Tags from the Samuel Roberts Noble Foundation  
Medicago truncatula drought library  
Unpublished (2000)  
Contact: May GD  
Plant Biology Division  
The Samuel Roberts Noble Foundation  
2510 Sam Noble Parkway, Ardmore, OK 73402, USA  
Tel: 580 221 7391  
Fax: 580 221 7380  
Email: gdmay@noble.org  
Insert Length: 293 Std Error: 0.00  
Plate: 082 row: F column: 04  
Seq primer: TCACACAGCAACACGCTATACAC.  
location/Qualifiers

FEATURES  
source  
1.293  
/organism="Medicago truncatula"  
/db\_xref="taxon:3880"

/clone="NF082F04DT"  
/clone\_lib="Drought"  
/issue\_type="Plantlets"  
/dev\_stage="Pooled timepoints"  
/note="Vector: Lambda Zap; Contains a mixture of entire  
plantlets harvested in a series of days post-watering  
timepoints."

BASE COUNT 109 a 76 c 25 g 83 t  
ORIGIN

Query Match 63.6%; Score 14; DB 11; Length 293;  
Best Local Similarity 63.6%; Pred. No. 2.2e+03;  
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Oy 1 tggtagannnnnnnttgata 22  
|||||||  
Db 119 tcttgcaatttcctcatttgata 98

RESULT 11  
AA515145 298 bp mRNA EST 20-AUG-1997  
LOCUS AA515145  
DEFINITION ng68c01.s1 NCI-CGAP\_Lip2 Homo sapiens cDNA clone IMAGE:93936  
similar to contains Alu repetitive element; contains element MSX1  
repetitive element; , mRNA sequence.  
ACCESSION AA515145  
VERSION AA515145.1 GI:2254745  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 298)  
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapsb@mail.nih.gov  
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: David B. Krizman, Ph.D.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution Information can be  
found through the I.M.A.G.E. Consortium/BLN at:  
www-bio.linn.gov/dbp/image/image.html  
Seq primer: -40ml3 fwd. ET from Amersham  
High quality sequence stop: 295.  
Location/Qualifiers

FEATURES  
source  
1.298  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:93936"  
/clone\_lib="NCI-CGAP\_Lip2"  
/issue\_type="liposarcoma"  
/lab\_host="DH10B"  
/note="Vector: pAMP10; mRNA made from liposarcoma, cDNA  
made by oligo-dT priming. Non-directionally cloned.  
Size-selected on agarose gel, average insert size 600 bp.  
Reference: Krizman et al. (1996) Cancer Research  
56:5380-5383."

BASE COUNT 76 a 63 c 91 g 68 t  
ORIGIN

Query Match 63.6%; Score 14; DB 10; Length 298;  
Best Local Similarity 63.6%; Pred. No. 2.2e+03;  
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Oy 1 tggtagannnnnnnttgata 22  
|||||||

Db 5 TGTGTGAGATGACATTTGATA 26

RESULT 12

LOCUS BE202756 332 bp mRNA EST 07-SEP-2000

DEFINITION EST402778 KVI Medicago truncatula cDNA clone pKV1-3023, mRNA sequence.

ACCESSION BE202756

VERSION BE202756.1 GI:8746021

KEYWORDS EST.

ORGANISM barrel medic.  
Medicago truncatula  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae; Medicago.

REFERENCE 1 (bases 1 to 332)  
Vandenbosch,K., Endre,G., Hur,J., Moore,J., Beremand,P., Ellis,L., Town,C.D., Bowman,C.L., Craven,M.B., Hansen,T.S., Holt,I.E. and Fraser,C.M.  
ESTs from roots of Medicago truncatula 24 hours after inoculation with Sinorhizobium meliloti  
Unpublished (1999)  
Contact: Vandenbosch K  
Department of Biology  
Texas A&M University  
College Station, TX 77843-3258, USA  
Tel: 409 845 7707  
Fax: 409 845 2891  
Email: kate@mail.bio.tamu.edu  
Texas A&M University:T262522e  
TIGR sequence name:MTIA966rk  
More information is available at:  
<http://chryslie.tamu.edu/medicago>  
Seq primer: SKmd (CTA gAA CTA gTg gAT CC).  
Location/Qualifiers

FEATURES

source

1..332  
/organism="Medicago truncatula"  
/cultivar="genotype A17"  
/db\_xref="taxon:3880"  
/clone="pKV1-3023"  
/clone\_lib="KVI"  
/tissue\_type="Seedling roots"  
/dev\_stage="24 hours post-inoculation with Sinorhizobium meliloti"  
/lab\_host="E. coli strain XL0LR"  
/note="Vector: pBluescript SK-; Site\_1: EcoRI; Site\_2: XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA was directionally ligated into the unique XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-Zap phage using Ex-assist helper phage and propagated in XL0LR cells."

BASE COUNT 95 a 50 c 96 g 91 t

ORIGIN

Query Match 63.6%; Score 14; DB 10; Length 332;  
Best Local Similarity 63.6%; Pred. No. 2.2e+03;  
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Oy 1 tgttgagannnnnnnnttgata 22  
||||||| |||||||

Db 157 TGTGTGATGATGCGTTTGATA 178

RESULT 13

LOCUS AZ037213 336 bp DNA GSS 01-MAR-2000

DEFINITION RPCI-23-364L15.TV RPCI-23 Mus musculus genomic clone RPCI-23-364L15, DNA sequence.

ACCESSION AZ037213

VERSION AZ037213.1 GI:7125453

KEYWORDS GSS.

SOURCE house mouse.

ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 336)  
Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akintet,B., Levins,M., McGann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P. and Fraser,C.M.  
Mouse BAC End Sequences from Library RPCI-23  
Unpublished (1999)  
Other\_GSSs: RPCI-23-364L15.TV  
Contact: Shaying Zhao  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
7712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: szhao@tigr.org  
Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.bufileo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.bufileo.edu/orderingframe.htm>) or from Resea ch Genetics (info@resgen.com). BAC end page: [http://www.tigr.org/tdb/bac\\_ends/mouse/bac\\_end\\_intro.html](http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html)  
Plate: 364 row: L column: 15  
Seq primer: T7  
Class: BAC ends.  
Location/Qualifiers

FEATURES

source

1..336  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="RPCI-23-364L15"  
/clone\_lib="RPCI-23"  
/sex="Female"  
/lab\_host="DH10B"  
/note="Organ: Kidney/Brain; Vector: pBAC3.6; Site\_1: EcoRI; Site\_2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBAC3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."

BASE COUNT 97 a 37 c 45 g 157 t

ORIGIN

Query Match 63.6%; Score 14; DB 13; Length 336;  
Best Local Similarity 63.6%; Pred. No. 2.2e+03;  
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Oy 1 tgttgagannnnnnnnttgata 22  
||||||| |||||||

Db 268 TGTGTGATTTCTCTTTTGATA 269

RESULT 14

LOCUS BH083614/C 342 bp DNA GSS 18-JUL-2001

DEFINITION RPCI-24-293E7.TJ RPCI-24 Mus musculus genomic clone RPCI-24-293E7, DNA sequence.

ACCESSION BH083614

VERSION BH083614.1 GI:14903211

KEYWORDS GSS.

ORGANISM house mouse.  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 342)  
Zhao,S., Nierman,W., Malek,J., Shatsman,S., Akintet,B., Levins,M., Tsegaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregorgis,E.,

TITLE Russell, D., de Jong, P. and Fraser, C.M.  
JOURNAL Mouse BAC End Sequences from Library RPCI-24  
COMMENT Unpublished (1998)

Contact: Shaying Zhao  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208

Email: szhao@tigr.org  
Clones are derived from the mouse BAC library RPCI-24. For BAC library availability, please contact Pieter de Jong (pdejong@tigr.org). Clones may be purchased from BACPAC Resources (<http://www.chori.org/bacpac/orderingframe.htm>). BAC end page: [http://www.tigr.org/tldb/bac\\_ends/mouse/bac\\_end\\_intro.html](http://www.tigr.org/tldb/bac_ends/mouse/bac_end_intro.html)  
Plate: 293 row: E column: 7  
Seq primer: SP6  
Class: BAC ends.

FEATURES  
source Location/Qualifiers

1..342  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="RPCI-24-293E7"  
/clone\_lib="RPCI-24"  
/sex="Male"  
/cell\_type="Spleen/Brain"  
/note="Vector: pTARBAC1; Site 1: BamHI; Site 2: BamHI; RPCI-24 Mouse BAC library produced by Pieter de Jong. The library was cloned in the pTARBAC1 cloning vector at the BamHI sites using MboI partially digested male C57BL/6J DNA."

BASE COUNT 99 a 57 c 87 g 99 t  
ORIGIN

Query Match 63.6%; Score 14; DB 13; Length 342;  
Best Local Similarity 63.6%; Pred. No. 2.2e+03;  
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 1 tgttgannnnnnnttgata 22  
|||||  
Db 158 TGTGGACCTTTGCTTGATA 137

## RESULT 15

AU112918 349 bp mRNA EST 19-OCT-2000  
LOCUS AU112918 unpublished oligo-capped cDNA library Caenorhabditis  
DEFINITION elegans cDNA YK7486 5', mRNA sequence.  
ACCESSION AU112918  
VERSION AU112918.1 GI:10926485  
KEYWORDS EST.

SOURCE Caenorhabditis elegans.

ORGANISM Caenorhabditis elegans.  
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditiida; Rhabditoidea  
; Rhabditiidae; Peloderinae; Caenorhabditis.

REFERENCE 1 (bases 1 to 349)  
AUTHORS Kohara, Y., Shin-I, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y.  
and Sugano, S.  
A complementary view of the C. elegans genome  
Unpublished (2000)

TITLE A complementary view of the C. elegans genome  
JOURNAL Unpublished (2000)  
COMMENT Contact: Yuji Kohara  
Genome Biology Lab.  
National Institute of Genetics  
Yata 1111, Mishima, Shizuoka 411, Japan  
Tel: 81-559-81-6854  
Fax: 81-559-81-6855  
Email: ykohara@lab.nig.ac.jp.

FEATURES  
source Location/Qualifiers

1..349  
/organism="Caenorhabditis elegans"  
/strain="N2"

/db\_xref="taxon:6239"  
/clone="YK74866"  
/clone\_lib="unpublished oligo-capped cDNA library"  
/sex="Hermaphrodite"  
/tissue\_type="whole animal"  
/dev\_stage="varied"

BASE COUNT 98 a 55 c 81 g 115 t  
ORIGIN

Query Match 63.6%; Score 14; DB 10; Length 349;  
Best Local Similarity 63.6%; Pred. No. 2.2e+03;  
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 tgttgannnnnnnttgata 22  
|||||  
Db 279 TGTGGATATCGATTTTGTGATA 300

Search completed: December 15, 2001, 02:33:48  
Job time: 4988 sec



**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 03:19:43 ; Search time 2725.73 Seconds  
(without alignments)  
133.153 Million cell updates/sec

Title: US-09-380-826A-2

Perfect score: 22  
Sequence: 1 tttgtgacacagattgata 22

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 1472140 segs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sy:\*  
12: gb\_un:\*  
13: gb\_vl:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_om:\*  
20: em\_or:\*  
21: em\_ov:\*  
22: em\_pat:\*  
23: em\_ph:\*  
24: em\_pl:\*  
25: em\_ro:\*  
26: em\_sy:\*  
27: em\_un:\*  
28: em\_vl:\*  
29: em\_vl:\*  
30: em\_htgo\_hum:\*  
31: em\_htgo\_inv:\*  
32: em\_htgo\_rnd:\*  
33: em\_htg\_hum:\*  
34: em\_htg\_inv:\*  
35: em\_htg\_rnd:\*  
36: em\_htg\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	22	100.0	1481	1	LFU60594	U60594 Leptospira
2	20.4	92.7	354	1	AB007014	AB007014 Spirochaeta
3	18.8	85.5	283	1	AF003953	AF003953 Leptospira
4	18.8	85.5	288	1	LIU94974	LIU94974 Leptospira
5	18.8	85.5	288	1	LIU94975	LIU94975 Leptospira
6	18.8	85.5	288	1	LIU94976	LIU94976 Leptospira
7	18.8	85.5	288	1	LIU94977	LIU94977 Leptospira
8	18.8	85.5	288	1	LIU94978	LIU94978 Leptospira
9	18.8	85.5	288	1	LIU94979	LIU94979 Leptospira
10	18.8	85.5	353	1	AB007012	AB007012 Spirochaeta
11	18.8	85.5	1012	8	ATHRNS2X	ATHRNS2X
12	18.8	85.5	9423	8	AT725979	AT725979 Arabidopsis
13	18.8	85.5	72590	8	AC003000	AC003000 Arabidopsis
14	18.8	85.5	121524	2	AF165146	AF165146 Homo sapi
15	18.8	85.5	129090	2	AC083961	AC083961 Homo sapi
16	18.8	85.5	182545	2	AC046176	AC046176 Homo sapi
17	18.8	85.5	200799	8	AF137379	AF137379 Nephrosol
18	18.8	85.5	200799	8	AF137379	AF137379 Nephrosol
19	18.4	83.6	1812	3	AF220067	AF220067 Drosophila
20	18.4	83.6	27365	2	AC015209	AC015209 Drosophila
21	18.4	83.6	173613	3	AC007475	AC007475 Drosophila
22	18.4	83.6	262731	3	AE003823	AE003823 Drosophila
23	17.8	80.9	1874	14	EBHPC44NCP	EBHPC44NCP
24	17.8	80.9	31214	8	SPCC830	SPCC830
25	17.8	80.9	85992	8	AB009052	AB009052 Arabidopsis
26	17.8	80.9	91894	2	AC026327	AC026327 Homo sapi
27	17.8	80.9	172681	9	AC026320	AC026320 Homo sapi
28	17.8	80.9	270889	2	AC055742	AC055742 Homo sapi
29	17.4	79.1	59261	8	T12M4	T12M4
30	17.4	79.1	81662	8	AB008265	AB008265 Arabidopsis
31	17.4	79.1	109476	9	AL513548	AL513548 Human DNA
32	17.4	79.1	109512	2	AC068144	AC068144 Homo sapi
33	17.4	79.1	148997	9	AC021015	AC021015 Homo sapi
34	17.4	79.1	187847	2	AC023155	AC023155 Homo sapi
35	17.2	78.2	316	1	AF352068	AF352068 Peptostre
36	17.2	78.2	462	1	AF175672	AF175672 unculture
37	17.2	78.2	835	8	AF271231	AF271231 Albino ca
38	17.2	78.2	1083	8	AF018567	AF018567 Unidentif
39	17.2	78.2	2213	8	LEP4CCOALB	LEP4CCOALB
40	17.2	78.2	3608	3	SCYLLO31C	SCYLLO31C
41	17.2	78.2	5596	8	SPU40832	SPU40832 Strongyloce
42	17.2	78.2	34496	3	U41016	U41016 Caenorhabdi
43	17.2	78.2	38655	3	U41545	U41545 Caenorhabdi
44	17.2	78.2	49311	2	F26K10	AL049803 Arabidops
45	17.2	78.2	60406	9	AL590550	AL590550 Human DNA

## ALIGNMENTS

RESULT	1	LOCUS	LFU60594	1481 bp	DNA	BCI	10-SEP-1998
DEFINITION	Leptospira fainei 16S ribosomal RNA gene, partial sequence.						
ACCESSION	U60594						
VERSION	U60594.1	GI:1408219					
KEYWORDS							
SOURCE	Leptospira fainei.						
ORGANISM	Bacteria: Spirochaetales; Leptospiraceae; Leptospira.						
REFERENCE	1 (bases 1 to 1481)						
AUTHORS	Perolat,P., Chappel,R.J., Adler,B., Baranton,G., Bulach,D.M., Billinghurst,M.L., Letocart,M., Merlen,F. and Serrano,M.S.						
TITLE	Leptospira fainei sp. nov., isolated from pigs in Australia						
JOURNAL	Int. J. Syst. Bacteriol. 48 Pt 3, 851-858 (1998)						
MEDLINE	98404550						
REFERENCE	2 (bases 1 to 1481)						
AUTHORS	Adler,B., Chappel,R.J., Baranton,G., Bulach,D.M., Billinghurst,M.L., Letocart,M., Merlen,F., Serrano,M.S. and Perolat,P.						

TITLE Direct Submission  
JOURNAL Submitted (12-JUN-1996) Microbiology, Monash University, Wellington Rd., Clayton, VIC 3168, Australia

FEATURES  
Source  
Location/Qualifiers  
1. 1481  
/organism="Leptospira falnei"  
/strain="Hurstbridge"  
/db\_xref="taxon:48782"  
<1. >1481  
rRNA  
/product="16S ribosomal RNA"  
BASE COUNT 391 a 335 c 439 g 314 t 2 others  
ORIGIN

Query Match 100.0%; Score 22; DB 1; Length 1481;  
Best Local Similarity 100.0%; Pred. No. 2;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tgttgatcacagaattgata 22  
|||||  
Db 155 TGTGATCACAGATTGATA 176

RESULT 2  
AB007014 354 bp DNA BCT 13-OCT-1997  
LOCUS Spirochaeta sp. 16S rRNA gene, partial sequence.  
DEFINITION AB007014  
ACCESSION AB007014.1 GI:2516255  
VERSION 16S ribosomal RNA.  
KEYWORDS Spirochaeta sp. (sub\_species: Freshwater obligate oligotroph, strain: FO-95) DNA.  
SOURCE Spirochaeta sp.  
Bacteria; Spirochaetales; Spirochaetaceae; Spirochaeta.  
REFERENCE 1 (bases 1 to 354)  
Shin, M.-S.  
Direct Submission  
Submitted (05-SEP-1997) to the DDBJ/EMBL/Genbank databases. M1-Sun Shin, Laboratory of Marine Molecular Microbiology, Faculty of Agriculture, Kyoto University, Okwacho, Kitashirakawa, Sakyo-ku, Kyoto, Kyoto 606-01, Japan (E-mail: msun@kais.kyoto-u.ac.jp, Tel: 075-753-6224, Fax: 075-753-6226)  
2 (sites)  
Shin, M., Yoshinaga, T., Uchida, A. and Ishida, Y.  
Phylogenetic analysis by 16S rRNA gene sequencing of Obligate Oligotrophs Isolated from the northern basin of Lake Biwa (Mesotrophic Lake).  
JOURNAL Unpublished (1997)  
FEATURES  
Source  
Location/Qualifiers  
1. 354  
/organism="Spirochaeta sp."  
/strain="FO-95"  
/sub\_species="Freshwater obligate oligotroph"  
/db\_xref="taxon:28185"  
<1. >354  
rRNA  
/product="16S rRNA"  
BASE COUNT 91 a 79 c 111 g 73 t  
ORIGIN

Query Match 92.7%; Score 20.4; DB 1; Length 354;  
Best Local Similarity 95.5%; Pred. No. 13;  
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 tfttgatcacagaattgata 22  
|||||  
Db 133 TGTGATCACAGATCTGATA 154

RESULT 3  
AF003953 283 bp DNA BCT 30-MAY-1998  
LOCUS Leptospira interrogans strain 48/95 16S ribosomal RNA gene, partial  
DEFINITION

sequence.  
ACCESSION AF003953.1 GI:3169306  
VERSION AF003953.1  
KEYWORDS Leptospira interrogans.  
SOURCE Leptospira interrogans.  
ORGANISM Bacteria; Spirochaetales; Leptospiraceae; Leptospira.  
REFERENCE 1 (bases 1 to 283)  
Woo, T.H.S., Patel, B.K.C., Cinco, M., Smythe, L.D., Symonds, M., Norris, M. and Dohnt, M.  
Identification of Leptoneima by real-time homogeneous assay of rapid cycle PCR product  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 283)  
Woo, T.H.S., Patel, B.K.C., Cinco, M., Smythe, L.D., Symonds, M., Norris, M. and Dohnt, M.  
Direct Submission  
Submitted (14-MAY-1997) School of Science, Griffith University, Brisbane, QLD 4111, Australia  
LOCATION/Qualifiers  
1. 283  
/organism="Leptospira interrogans"  
/strain="48/95"  
/db\_xref="taxon:173"  
<1. >283  
rRNA  
/product="16S rRNA"  
BASE COUNT 74 a 60 c 88 g 61 t  
ORIGIN

Query Match 85.5%; Score 18.8; DB 1; Length 283;  
Best Local Similarity 90.9%; Pred. No. 78;  
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 tgttgatcacagaattgata 22  
|||||  
Db 136 TATTGATCACAGATTGATA 157

RESULT 4  
U94974 288 bp DNA BCT 01-JAN-1998  
LOCUS Leptospira inadae 16S ribosomal RNA gene, partial sequence.  
DEFINITION U94974  
ACCESSION U94974.1 GI:2735446  
VERSION 16S ribosomal RNA.  
KEYWORDS Leptospira inadae.  
SOURCE Leptospira inadae.  
Bacteria; Spirochaetales; Leptospiraceae; Leptospira.  
REFERENCE 1 (bases 1 to 288)  
Woo, T.H.S., Smythe, L.D., Symonds, M., Norris, M., Dohnt, M., Brenner, D.J. and Patel, B.K.C.  
Identification of Leptospira inadae by continuously monitoring fluorescence during rapid cycle PCR  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 288)  
Woo, T.H.S., Smythe, L.D., Symonds, M., Norris, M., Dohnt, M., Brenner, D.J. and Patel, B.K.C.  
Direct Submission  
Submitted (24-MAR-1997) School of Science, Griffith University, Brisbane, QLD 4111, Australia  
LOCATION/Qualifiers  
1. 288  
/organism="Leptospira inadae"  
/strain="79/95"  
/db\_xref="taxon:29506"  
<1. >288  
rRNA  
/product="16S ribosomal RNA"  
BASE COUNT 77 a 61 c 88 g 62 t  
ORIGIN

Query Match 85.5%; Score 18.8; DB 1; Length 288;





TITLE	Brenner,D.J. and Patel,B.K.C. Identification of Leptospira inadai by continuous monitoring fluorescence during rapid cycle PCR
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 288)
AUTHORS	Woo,T.H.S., Smythe,L.D., Symonds,M., Norris,M., Dohnt,M., Brenner,D.J. and Patel,B.K.C.
TITLE	Direct Submission
JOURNAL	Submitted (24-MAR-1997) School of Science, Griffith University, Brisbane, Qld 4111, Australia
FEATURES	location/Qualifiers
source	1..288
	/organism="Leptospira inadai"
	/strain="218/95"
	/db_xref="taxon:29506"
trna	<1..>288
	/product="16S ribosomal RNA"
BASE COUNT	77 a 61 c 88 g 62 t
ORIGIN	

	Query Match	85.5%;	Score 18.8;	DB 1;	Length 288;
	Best Local Similarity	90.9%;	Pred. No. 78;		
	Matches 20;	Conservative	0;	Mismatches 2;	Indels 0;
					Gaps 0;
Oy	1	tggtgacacaaagattgata	22		
Db	141	TATTGGATCACAGATTTGATA	162		

RESULT		9				
LIIU94979						
LOCUS	LIIU94979	288 bp	DNA	BCT	01-JAN-1998	
DEFINITION	Leptospira inadai 16S ribosomal RNA gene, partial sequence.					
ACCESSION	U94979					
VERSION	U94979.1 GI:2735451					
KEYWORDS	.					
SOURCE	Leptospiara inadai.					
ORGANISM	Leptospiara inadai					

REFERENCE	1 (bases 1 to 288)
AUTHORS	Woo,T.H.S., Smythe,L.D., Symonds,M., Norris,M., Dohnt,M., Brenner,D.J. and Patel,B.K.C.
TITLE	Identification of <i>Leptospira</i> <i>inadai</i> by continuously monitoring fluorescence during rapid cycle PCR
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 288)
AUTHORS	Woo,T.H.S., Smythe,L.D., Symonds,M., Norris,M., Dohnt,M., Brenner,D.J. and Patel,B.K.C.
TITLE	Direct Submission
JOURNAL	Submitted (24-MAR-1997) School of Science, Griffith University
FEATURES	Brisbane, QLD 4111, Australia Location/Qualifiers

```

ORIGIN      77 a      61 c      88 g      62 t
BASE COUNT
          T RNA
          <1. ->268 /Product="16S ribosomal RNA"
          /Db_xref="Ezaxon:29506"
          /strain="1078 VR1"
          /organism="Neopospira inadai"

```

Query Match	85.58%	Score 18.8;	DB 1;	Length 288;
Best Local Similarity	90.98%	Pred. No. 78;		
Matches 20; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0

```
QY 1 tgttgatcacaaattgata 22
    | | | | | | | | | | | |
Db 141 TATTGCATCACAGATTGATA 162
```

## RESULT 10

AB007012	AB007012	353 bp	DNA	. ECT	13-OCT-1997
LOCUS	AB007012	353 bp	DNA	. ECT	13-OCT-1997
DEFINITION	Spirocheta sp. 16S rRNA gene, partial sequence.				
ACCESSION	AB007012				
VERSION	AB007012.1	GI:2516253			
KEYWORDS	16S ribosomal RNA				
SOURCE	Spirocheta sp. sub-species: Freshwater obligate oligotroph,				

ORGANISM  
Spirochaeta sp.  
Bacteria: Spirochaetales; Spirochaetaceae; Spirochaeta.  
REFERENCE  
1 (bases 1 to 355)  
AUTHORS  
Shin,M.-S.  
TITLE  
Direct Submission  
JOURNAL  
Submitted (05-SEP-1997) to the DDBJ/EMBL/GenBank databases. M1-Sun-  
Submitted (05-SEP-1997) to the DDBJ/EMBL/GenBank databases. M1-Sun-  
Submitted (05-SEP-1997) to the DDBJ/EMBL/GenBank databases. M1-Sun-

REFERENCE	2 (sites)
AUTHORS	Shin, M., Yoshinaga, I., Uchida, A. and Ishida, Y.
TITLE	Phylogenetic analysis by 16S rRNA gene sequencing of <i>Obligate</i> <i>Oligotrophs</i> isolated from the northern basin of Lake Biwa (Mesotrophic Lake)
JOURNAL	unpublished (1997)

JOURNAL	Unpublished (1991)
FEATURES	Location/Qualifiers
source	1. .353

```

/sub_species="Freshwater obligate oligotroph"
/db_xref="taxon:28185"
<1.1>353
rRNA

```

	/product="165 IRNA"			
BASE COUNT	90 a	80 c	108 g	75 t
ORIGIN				

Query Match	85.5%	Score 18.8	DB 1	length 353
Best Local Similarity	90.9%	Pred. No. 77		
Matches 20	Conservative	0	Mismatches 2	Indels 0
				Gaps 0

```

QY      1  tgttgatcacagaattgata  22
          |||||
Db      132 TGTGGATCACCAGATCTGATA  153

```

RESULT	11
ATHRSNS2X	
LOCUS	1012 bp mRNA
DEFINITION	Aarabidopsis thaliana ribonuclease (RNS2) mRNA, complete cds.
ACCESSION	M68336
VERSION	M68336.1 GI:289209
KEYWORDS	ribonuclease.
SOURCE	Arabidopsis thaliana cDNA to mRNA.

REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots: Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 1012)
AUTHORS	Taylor,C.B., Bartola,P.A., delcardayre,S.B., Ratnes,R.T. and Green,P.J.
TITLE	Rn2: a senescence-associated RNase of Arabidopsis that diverged from the S-RNases before speciation
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 90 (11), 5118-5122 (1993)
MEDLINE	
FEATURES	Location/Qualifiers

Source

gene

CDS

/codon\_start=1

		/product="ribonuclease"	
		/protein_id="AA51406.1"	
		/db_xref="GI:289210"	
		/translation="MASRLCLLLVACIAGAFADVI ELNRSQREDFYALSLQPCT	
		YCRTRHCCSKNMACCRCGSDAPTOFTIGLWPDVNDGSPSCYSRDKKEITSLMGGL	
		LEKWPFLSCGSPSSCGNGCKSGFWGHMEKGTCSVPFDEYNYFLTTLNLTKHNW	
		TDVLYOAGVYASNSSEKYPLGIIVAIIONAETFEVVKRAIDIRICEFYKDKPRDC	
		CVGSDLTLSRKSCPKYVSLEPYPLDDEAMVLKMPFEREAL"	
BASE COUNT	287 a	197 c	217 g 311 t
ORIGIN			
Query Match	85.5%	Score 18.8;	DB 8; length 1012;
Best Local Similarity	90.9%;	Ared. No. 76;	
Matches 20;	Conservative 0;	Mismatches 2;	Indels 0; Gaps 0;
Oy	1	tgttgatcacagaattgata 22	
Db	672	TGTTGGTTCACAGAATTGACA 693	
RESULT 12			
LOCUS	ATH275979	9423 bp DNA	PLN 29-MAR-2001
DEFINITION	Arabidopsis thaliana gene for GDP-mannose pyrophosphorylase, gene for vacuolar ribonuclease and gene for proline biosynthesis codifying enzyme.		
ACCESSION	AJ275979		
VERSION	AJ275979.1 GI:13509286		
KEYWORDS	GDP-mannose pyrophosphorylase; proline biosynthesis codifying enzyme; vacuolar ribonuclease.		
SOURCE	thale cress.		
ORGANISM	Arabidopsis thaliana		
	Eumariophyta: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots: Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids. 1 (bases 1 to 9423)		
REFERENCE	Genes responding to phosphate starvation placed together in Arabidopsis genome		
AUTHORS	Unpublished		
TITLE	2 (bases 1 to 9423)		
JOURNAL	Direct Submission		
REFERENCE	Submitted (01-MAR-2000) AVILA C., Biologia Molecular y Bioquímica, Facultad de Ciencias, E-29071, Málaga, SPAIN		
JOURNAL	Location/Qualifiers		
FEATURES	1..9423 /organism="Arabidopsis thaliana"		
source	/db_xref="taxon:3702"		
	1405..1515		
exon	/number=1		
	join(1405..1515,1597..1659,1748..1856,1883..2013,2099..2270)		
CDS	/codon_start=1		
	/product="GDP-mannose pyrophosphorylase"		
	/protein_id="CAC35355.1"		
exon	/db_xref="GI:13509286"		
	/translation="MKALLLVGGFGTRLRPILTSFPPELVDFANKPMILHOIALKANVGVDEVDIAINQPEVMENLFKLDFKTLEIKTCQEQEPYLGTAAPLARLDKGSGGEPEPVLNSDVISEYPLKEILEFHFKSHGGEASIMTKVDEPSKYVVAMEESTGVKEKVEVERKLVGNKINAGIYLILANSVDKIELRPTSELKEPEPIIAAOOGVAVLGPFWMMDIQGRDYINGCIRLYLDSIRKKSPAKLISGPHLYGANGVINDPETAIGSCILGPRVAIMGPGCVESGVRLSRCTYMRGVRIKKKHACISSITICWHSTVYGOMARIENNNTIIIGEYHVHSDELISNGGVLPHEIKSNILKPEIWM		
intron	join(1405..1515,1597..1659,1748..1856,1883..2013,2099..2270)		
exon	/number=1		
	1516..1596		
intron	/number=1		
	1597..1659		
exon	/number=2		
	1660..1747		
Intron	/number=2		

exon		1748..1856	/number=3
intron		1857..1882	/number=3
exon		1883..2013	/number=4
intron		2014..2098	/number=4
exon		2099..2270	/number=5
CDS		3288..3473	/codon_start=1
		/product="vacuolar ribonuclease"	
		/protein_id="CAC35356.1"	
		/db_xref="GI:13509288"	
		/translation="MASRLCLLLLVACIAGAFADVIELNRSQREDFALSLQMPG	
		YCRGTRHCCSKMACRG"	
CDS		complement(8081..>8158)	
		/codon_start=1	
		/product="proline biosynthesis enzyme"	
		/protein_id="CAC35357.1"	
		/db_xref="GI:13509289"	
BASE COUNT	2795 a	1693 c	1760 g 3155 t
ORIGIN			
Query Match	85.5%:	Score 18.8;	DB 8; Length 9423;
Best Local Similarity	90.9%:	Pred. No. 74;	
Matches	20; Conservative	0; Mismatches	2; Indels 0; Gaps 0
OY	1	tgttgatccacaagattgata 22	
Db	5307	TGTTGGTTCACAGATTGTACA	5328
RESULT_13			
LOCUS	AC003000	72590 bp	DNA PLN 05-APR-2000
DEFINITION		Arabidopsis thaliana chromosome II section 214 of 255 of the complete sequence. Sequence from clones T517.	
ACCESSION	AC003000	AEO02093	
VERSION	AC003000.2	GI:6598383	
KEYWORDS		HFG.	
SOURCE		thale cress.	
ORGANISM		Arabidopsis thaliana	
REFERENCE		Eukaryota: Viridiplantae: Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots: Rosidae: euroids II; Brassicales; Brassicaceae; Arabidopsids. 1 (bases 1 to 72590)	
AUTHORS		Liu,X., Kaul,S., Rounsley,S.D., Shea,T.P., Benito,M.-I., Town,C.D., Fujii,C.Y., Mason,T.M., Bowman,C.L., Barnstead,M.E., Feldlyum,T.V., Buell,C.R., Ketchum,K.A., Lee,J.J., Ronning,C.M., Koo,H., Moffat,K.S., Cronin,L.A., Shen,M., Vanden,S.E., Umayam,L., Tallon,L.J., Gill,J.E., Adams,M.D., Carrera,A.J., Creasy,T.H., Goodman,H.M., Somerville,C.R., Copenhagen,G.P., Preuss,D., Nierman,W.C., White,O., Eisen,J.A., Salzberg,S.L., Fraser,C.M. and Venter,J.C.	
TITLE		Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana	
JOURNAL		Nature 402 (6763), 761-768 (1999)	
MEDLINE		20083487	
PUBMED		10617197	
REFERENCE		2 (bases 1 to 72590)	
AUTHORS		Liu,X.	
TITLE		Direct Submission	
JOURNAL		Submitted (09-MAR-2000) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA	
COMMENT		On Dec 17, 1999 this sequence version replaced gi:2642152. The sequence and annotation of chromosome 2 were merged from those of the individual clones on this chromosome after removing overlaps. For detailed information, please see the TIGR web site ( <a href="http://www.tigr.org/tcdb/at/at.html">http://www.tigr.org/tcdb/at/at.html</a> ).	



```
/gene="At2g39740"
/feature="hypothetical protein"
/codon_start=1
/db_xref="GI:2642156"
/translation="MSRNPFLDPTLOEILQVIKPRADRDRTITVIDLRDLYQSEV
LKGAVQPFQSVSNLFTRMGDLDISVDFSGSILFTGKKOKQILLHLALRASG
LWKLQFVHARVPLIKVSGHORISIDNLDGLKSRFLVISEIDGFRDLVL
LVKEAKAHNINDSKTGFNSLSLVIFHRCVPAVILPELVIPKSAVDLTV
RKTAEESIAOVTANIANIARFKSRASVNRSSISELVSEFFAKVEPPEQPVNAASV
RNRDRIADVFOITSRLVSECNRSIIIGLGHIOESLVFTISLPSQIHANGMHV
RNLHQARPQNDQMOQNSQSYNTPNPMPPLTSPQOQNTQNNPRLQGOQVQO
QTPVITQTOQOKSPKSGNRPKNTSAGSSQNGHIGKPSGHNHNVASAPATYNG
VNSARPPSKIPSGOIVMRPRHEQ"
complement(join(21380..21726,21839..21916,22124..22246,
22328..22585,22834..22932,23212..23322,23481..23751,
23847..23996,24089..24814))
/gene="At2g39750"
complement(21380..24814)
```

Query Match  
Best Local Similarity 85.5%; Score 18.8; DB 8; Length 72590;  
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 tgttgatcacagaattgata 22  
|||||  
Db 35807 TGTGTGTCACAAGATTGTGACA 35828

RESULT 14  
AF165146 121524 bp DNA HTG 07-JUN-2001  
LOCUS Homo sapiens chromosome 8 clone CTA-397H3 map 8q12-8q13, \*\*\*  
DEFINITION SEQUENCING IN PROGRESS \*\*\*, 7 unordered pieces.  
ACCESSION AF165146  
VERSION AF165146.3 GI:14327840  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_FULTOP.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 121524)  
Schlilabel,M.B., Baumgart,C., Blechschmidt,K., Dete,M., Jahn,N.,  
Lehmann,R., Menzel,U., Polley,A., Reichwald,K., Schudy,A.,  
Siddiqui,R., Taudien,S., Wen,G., Schlegelberger,B., Siebert,R.,  
Rosenthal,A. and Platzer,M.  
Chromosome 8 genomic sequence  
Unpublished  
2 (bases 1 to 121524)  
Schudy,A., Blechschmidt,K., Schlilabel,M., Baumgart,C., Menzel,U.,  
Weber,U., Schatchevoy,R. and Rosenthal,A.  
Direct Submission  
Submitted (06-JUN-1999) Genome Analysis, Institute of Molecular  
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany  
On Jun 7, 2001 this sequence version replaced g1:8151945.  
----- Genome Center  
Center: Institute of Molecular Biotechnology  
Center code: IMB  
Web site: http://genome.imb-jena.de/  
Contact: gscj-submit@genome.imb-jena.de  
----- Project Information  
Center project name: H171  
Center clone name: CTA-397H3  
----- Summary Statistics  
Sequencing vector: M13; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; Version 0.990329  
Consensus quality: 115981 bases at least Q40  
Consensus quality: 117981 bases at least Q30  
Consensus quality: 119332 bases at least Q20  
Quality coverage: 9.57 in Q20 bases; sum-of-contigs

\* NOTE: This is a 'working draft' sequence. It currently

\* consists of 7 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
1 1187: contig of 1187 bp in length  
1 1188 1287: gap of unknown length  
\* 1188 3530: gap of 2243 bp in length  
\* 1288 3630: gap of unknown length  
\* 3531 21624: contig of 17994 bp in length  
\* 3631 21724: gap of unknown length  
\* 21625 43656: contig of 21932 bp in length  
\* 21725 43657 43756: gap of unknown length  
\* 43657 61383: contig of 17627 bp in length  
\* 43757 61384 61484: gap of unknown length  
\* 61384 85418: contig of 23835 bp in length  
\* 61484 85418: gap of unknown length  
\* 85319 121524: contig of 36106 bp in length.  
\* 85419  
Location/Qualifiers  
1.121524  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="8"  
/map="8q12-8q13"  
/clone="CTA-397H3"

BASE COUNT 34959 a 25555 c 25579 g 34831 t 600 others  
ORIGIN

Query Match  
Best Local Similarity 85.5%; Score 18.8; DB 2; Length 121524;  
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 tgttgatcacagaattgata 22  
|||||  
Db 7103 TTTTGATCACAAGATTGTGATA 7124

RESULT 15  
AC083961  
LOCUS Homo sapiens chromosome 8 clone RP11-22E14 map 8, \*\*\* SEQUENCING IN  
DEFINITION PROGRESS \*\*\*, 67 unordered pieces.  
ACCESSION AC083961  
VERSION AC083961.2 GI:12831382  
KEYWORDS HTG; HTGS\_PHASE1.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 129090)  
Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
Unpublished  
2 (bases 1 to 129090)  
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,  
Anderson,S., Barna,N., Bastien,V., Beda,F., Boguslavsky,L.,  
Bouhagalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,  
Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,P.,  
Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,  
Fitzhugh,W., Gage,D., Galagan,J., Gardina,S., Glade,S., Goyette,M.,  
Graham,L., Grand-pierre,N., Hagos,B., Heaford,A., Horton,L.,  
Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., LaRocque,K.,  
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G.,  
Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McKernan,K.,  
McPheeters,R., Meldrim,D., Meneses,L., Mihova,T., Mlenga,V.,  
Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T.,  
O'Donnell,P., O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K.,  
Pierre,N., Pisanl,C., Polara,V., Raymond,C., Riebeck,M., Riley,R.,  
Rogov,P., Rotman,D., Roy,A., Santos,R., Schauer,S., Severy,P.,  
Sougnuez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,

TITLE  
JOURNAL  
COMMENT

Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodores, J.,  
Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A.,  
Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zahoun, J.,  
Zimmer, A., and Zody, M.  
Submitted (08-OCT-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Feb 15, 2001 this sequence version replaced gi:10717226.  
All repeats were identified using RepeatMasker:  
Smith, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

----- Project Information

Center Project name: L11257

Center Clone name: 22\_E\_14

-----

\* NOTE: This is a 'working draft' sequence. It currently

\* consists of 67 contigs. The true order of the pieces

\* is not known and their order in this sequence record is

\* arbitrary. Gaps between the contigs are represented as

\* runs of N, but the exact sizes of the gaps are unknown.

\* This record will be updated with the finished sequence

\* as soon as it is available and the accession number will

\* be preserved.

\* 1 1369: contig of 1369 bp in length

\* 1370 1469: gap of 100 bp

\* 1470 2279: contig of 810 bp in length

\* 2280 2379: gap of 100 bp

\* 2380 3192: contig of 813 bp in length

\* 3193 3292: gap of 100 bp

\* 3293 3539: contig of 247 bp in length

\* 3540 3639: gap of 100 bp

\* 3640 4530: contig of 891 bp in length

\* 4531 4630: gap of 100 bp

\* 4631 5236: contig of 606 bp in length

\* 5237 5336: gap of 100 bp

\* 5337 6055: contig of 719 bp in length

\* 6056 6155: gap of 100 bp

\* 6156 6653: contig of 498 bp in length

\* 6654 6753: gap of 100 bp

\* 6754 7388: contig of 635 bp in length

\* 7389 7488: gap of 100 bp

\* 7489 8145: contig of 657 bp in length

\* 8146 8245: gap of 100 bp

\* 8246 9151: contig of 906 bp in length

\* 9152 9251: gap of 100 bp

\* 9252 9962: contig of 711 bp in length

\* 9963 10062: gap of 100 bp

\* 10063 10776: contig of 714 bp in length

\* 10777 10876: gap of 100 bp

\* 10877 11474: contig of 598 bp in length

\* 11475 11574: gap of 100 bp

\* 11575 12639: contig of 1065 bp in length

\* 12640 12739: gap of 100 bp

\* 12740 13952: contig of 1213 bp in length

\* 13953 14052: gap of 100 bp

\* 14053 14638: contig of 786 bp in length

\* 14639 14938: gap of 100 bp

\* 14939 15849: contig of 911 bp in length

\* 15850 15949: gap of 100 bp

\* 15950 18036: contig of 2087 bp in length

\* 18037 18136: gap of 100 bp

\* 18137 19282: contig of 1146 bp in length

\* 19283 19382: gap of 100 bp

\* 19383 20437: contig of 1055 bp in length

\* 20438 20537: gap of 100 bp

\* 20538 21159: contig of 622 bp in length

\* 21160 21259: gap of 100 bp

\* 21260 22494: contig of 1235 bp in length

\* 22495 22594: gap of 100 bp

\* 22595 23478: contig of 884 bp in length

\* 23479 23578: gap of 100 bp

\* 23579 24600: contig of 1022 bp in length

\* 24601 24700: gap of 100 bp

\* 24701 24891: contig of 191 bp in length

\* 24892 24991: gap of 100 bp

\* 24992 26183: contig of 1192 bp in length

\* 26184 26283: gap of 100 bp

\* 26284 27862: contig of 1579 bp in length

\* 27863 27962: gap of 100 bp

\* 27963 28581: contig of 619 bp in length

\* 28582 28681: gap of 100 bp

\* 28682 29511: contig of 830 bp in length

\* 29512 29611: gap of 100 bp

\* 29612 31038: contig of 1427 bp in length

\* 31039 31138: gap of 100 bp

\* 31139 32562: contig of 1444 bp in length

\* 32563 32682: gap of 100 bp

\* 32683 33540: contig of 858 bp in length

\* 33541 33640: gap of 100 bp

\* 33641 35263: contig of 1623 bp in length

\* 35264 35363: gap of 100 bp

\* 35364 37036: contig of 1673 bp in length

\* 37037 37136: gap of 100 bp

\* 37137 38965: contig of 1827 bp in length

\* 38964 39063: gap of 100 bp

\* 39064 40807: contig of 1744 bp in length

\* 40808 40907: gap of 100 bp

\* 40908 42896: contig of 1989 bp in length

\* 42897 42996: gap of 100 bp

\* 42997 44598: contig of 1602 bp in length

\* 44599 44698: gap of 100 bp

\* 44699 46343: contig of 1645 bp in length

\* 46344 46443: gap of 100 bp

\* 46444 47574: contig of 1131 bp in length

\* 47575 47674: gap of 100 bp

\* 47675 50217: contig of 2543 bp in length

\* 50218 50317: gap of 100 bp

\* 50318 51338: contig of 1021 bp in length

\* 51339 51438: gap of 100 bp

\* 51439 53743: contig of 2305 bp in length

\* 53744 53843: gap of 100 bp

\* 53844 55260: contig of 1417 bp in length

\* 55261 55360: gap of 100 bp

\* 55361 57843: contig of 2483 bp in length

\* 57844 57943: gap of 100 bp

\* 57944 59865: contig of 1922 bp in length

\* 59866 59965: gap of 100 bp

\* 59966 62260: contig of 2295 bp in length

\* 62261 62360: gap of 100 bp

\* 62361 64120: contig of 1760 bp in length

\* 64121 64220: gap of 100 bp

\* 64221 65979: contig of 1759 bp in length

\* 65980 66079: gap of 100 bp

\* 66080 68712: contig of 2633 bp in length

\* 68713 68812: gap of 100 bp

\* 68813 71145: contig of 2233 bp in length

\* 71146 71245: gap of 100 bp

\* 71246 74155: contig of 2910 bp in length

\* 74156 74255: gap of 100 bp

\* 74256 77041: contig of 2786 bp in length

\* 77042 77141: gap of 100 bp

\* 77142 80685: contig of 3544 bp in length

\* 80686 80785: gap of 100 bp

\* 80786 83825: contig of 3040 bp in length

\* 83826 83925: gap of 100 bp

\* 83926 87721: contig of 3796 bp in length

\* 87722 87821: gap of 100 bp

\* 87822 92728: contig of 4607 bp in length

\* 92729 92828: gap of 100 bp

\* 92829 96652: contig of 3824 bp in length

\* 96653 96752: gap of 100 bp

```

* 96753 99906: contig of 3154 bp in length
* 99907 100006: gap of 100 bp
* 100007 102910: contig of 2904 bp in length
* 102911 103010: gap of 100 bp
* 103011 107188: contig of 4178 bp in length
* 107189 107288: gap of 100 bp
* 107289 111047: contig of 3759 bp in length
* 111048 111147: gap of 100 bp
* 111148 116741: contig of 5594 bp in length
* 116742 116841: gap of 100 bp
* 116842 121990: contig of 5149 bp in length
* 121991 122090: gap of 100 bp
* 122091 126622: contig of 4532 bp in length
* 126623 126722: gap of 100 bp
* 126723 129090: contig of 2368 bp in length.

```

# FEATURES

```

Location/Qualifiers
1. 129090
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8"
/clone="RP11-22E14"
/clone_lib="RPC1-11 Human Male BAC"

```

```

Query Match      85.5%; Score 18.8; DB 2; Length 129090;
Best Local Similarity 90.9%; Pred. No. 71;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 tgttgatcacagaattgata 22
    | | | | | | | | | | | | | |
Db 6294 TTTTGGATCACAAGATTGGTA 6315

```

Search completed: December 15, 2001, 03:19:51  
Job time: 7681 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 03:29:41 : Search time 401.91 Seconds  
(without alignments)  
46.929 Million cell updates/sec

Title: US-09-380-826a-2

Perfect score: 22  
Sequence: 1 tttgtatcacacagatttgata 22

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N.Geneseq\_1101:\*

1: /SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq/NA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	100.0	22	19	AAV58901
2	22	100.0	22	19	AAV58897
3	22	100.0	22	19	AAV58896
4	18.8	85.5	1012	21	AAV99365
5	18.8	85.5	1069	21	AAV37325
6	17.8	80.9	2084	21	AAV96966
7	17.2	78.2	447	21	AAV18114
8	17.2	78.2	1390	21	AAV42365
9	17.2	78.2	1413	21	AAV46258
10	17.2	78.2	2038	22	AAV22799
11	16.8	76.4	1764	22	AAV15352

C	12	16.8	76.4	2582	18	AAV90543	Maize male fertili
C	13	16.4	74.5	2444	21	AAV56059	Eucalyptus grandis
C	14	16.2	73.6	335	22	AAV33702	Human colon cancer
C	15	16.2	73.6	951	20	AAV84598	Human secreted pro
C	16	16.2	73.6	1250	21	AAV59054	Human secreted pro
C	17	16.2	73.6	1291	19	AAV59636	Human secreted pro
C	18	16.2	73.6	1376	22	AAV33007	Human colon cancer
C	19	16.2	73.6	1450	21	AAV49931	Arabidopsis thalia
C	20	16.2	73.6	1452	21	AAV39657	Arabidopsis thalia
C	21	16.2	73.6	1473	20	AAV61757	B. burgdorferi ant
C	22	16.2	73.6	2000	17	AAV17111	Acetyl-CoA-carboxy
C	23	16.2	73.6	2001	17	AAV39904	Maize acetyl CoA c
C	24	16.2	73.6	2001	17	AAV49816	ECORI fragment of
C	25	16.2	73.6	3822	18	AAV64683	M. leprae gyrA pre
C	26	16.2	73.6	4366	14	AAV42933	A3 maize ACCase CD
C	27	16.2	73.6	5400	21	AAV39732	Potato subclone pg
C	28	16.2	73.6	7470	17	AAV39905	Maize acetyl CoA c
C	29	16.2	73.6	7470	19	AAV29317	Maize ACCase enzym
C	30	16.2	73.6	7470	21	AAV49820	Maize acetyl CoA c
C	31	16.2	73.6	10811	19	AAV59091	Potato pollen cell
C	32	15.8	71.8	467	22	AAV32060	Human olfactory re
C	33	15.8	71.8	670	21	AAV09187	Human secreted pro
C	34	15.8	71.8	736	21	AAV10808	Human secreted pro
C	35	15.8	71.8	740	21	AAV51908	Arabidopsis thalia
C	36	15.8	71.8	843	20	AAV51567	Human gene express
C	37	15.8	71.8	961	19	AAV14095	H. pylori GHP0 131
C	38	15.8	71.8	4146	21	AAV51556	Human hypoxia resp
C	39	15.6	70.9	297	20	AAV88753	EST clone HK650.
C	40	15.6	70.9	342	22	AAV32175	Human olfactory re
C	41	15.6	70.9	452	22	AAV53651	S. epidermidis ope
C	42	15.6	70.9	635	22	AAV31988	Human olfactory re
C	43	15.6	70.9	795	20	AAV72025	Adenovirus pACT562
C	44	15.6	70.9	834	20	AAV72026	Adenovirus SCAR-RG
C	45	15.6	70.9	1075	18	AAV72715	C. elegans inhibit

#### ALIGNMENTS

RESULT	1
ID	AAV58901 standard; DNA: 22 BP.
XX	AAV58901;
AC	AAV58901;
XX	20-JAN-1999 (first entry)
DE	Leptospira rRNA gene nucleotide sequence.
XX	Infection; pathogenic Leptospira; protective immunity; therapy;
KW	diagnosis; ss.
XX	
OS	Leptospira sp.
XX	
PN	WO9840099-A1.
XX	
PD	17-SEP-1998.
XX	
PF	06-MAR-1998; 98WO-AU00145.
XX	
PR	07-MAR-1997; 97AU-0005494.
PA	(AGRI-) AGRIC VICTORIA SERVICES PTY LTD.
PA	(PIGR-) PIG RES & DEV CORP.
PI	Chappel RJ;
DR	WPI; 1998-520791/44.
PT	New isolated pathogenic Leptospira bacterium - useful for e.g
PT	developing products for conferring protective immunity, and for
PT	prophylactic or therapeutic treatment
XX	

PS Claim 15; Page 72; 94pp; English.  
XX  
CC This sequence represents a Leptospira DNA sequence isolated from the  
CC pathogenic Leptospira (LS) bacterium of the invention. The bacterium  
CC belongs to serogroup Hurstbridge of serovar hurstbridge or the species  
CC L. fainei. The LS bacteria can be used for conferring protective  
CC immunity against pathogenic LS bacteria in humans or animals. The  
CC bacteria can also be used for prophylactic or therapeutic treatment of LS  
CC infections. The DNAs and antibodies may also be used for detection and  
CC diagnosis of past or present LS infection.  
XX  
SQ Sequence 22 BP; 7 A; 2 C; 5 G; 8 T; 0 other;  
  
Query Match 100.0%; Score 22; DB 19; Length 22;  
Best Local Similarity 100.0%; Pred. No. 0.11; Mismatches 0; Indels 0; Gaps 0;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 1 tgttgatcacagaattgata 22  
|||||  
DB 1 tgttgatcacagaattgata 22  
  
RESULT 2  
AAV58897  
ID AAV58897 standard; DNA; 22 BP.  
XX  
AC AAV58897;  
XX  
DT 20-JAN-1999 (first entry);  
XX  
DE L. fainei nucleotide sequence.  
XX  
KW Infection; pathogenic Leptospira; protective immunity; therapy;  
KM diagnosis; ss.  
XX  
OS Leptospira fainei.  
XX  
PN WO9840099-A1.  
XX  
PD 17-SEP-1998.  
XX  
PE 06-MAR-1998; 98WO-AU00145.  
XX  
PR 07-MAR-1997; 97AU-0005494.  
XX  
PA (AGRI-) AGRIC VICTORIA SERVICES PTY LTD.  
XX (PIGR-) PIG RES & DEV CORP.  
XX  
PI Chappel RJ;  
XX  
DR WPI: 1998-520791/44.  
XX  
PT New isolated pathogenic Leptospira bacterium - useful for, e.g  
PT developing products for conferring protective immunity, and for  
PT prophylactic or therapeutic treatment  
XX  
PS Claim 15; Page 70; 94pp; English.  
XX  
CC This sequence represents a Leptospira DNA sequence isolated from the  
CC pathogenic Leptospira (LS) bacterium of the invention. The bacterium  
CC belongs to serogroup Hurstbridge or serovar hurstbridge or the species  
CC L. fainei. The LS bacteria can be used for conferring protective  
CC immunity against pathogenic LS bacteria in humans or animals. The  
CC bacteria can also be used for prophylactic or therapeutic treatment of LS  
CC infections. The DNAs and antibodies may also be used for detection and  
CC diagnosis of past or present LS infection.  
XX  
SQ Sequence 22 BP; 7 A; 2 C; 5 G; 8 T; 0 other;  
  
Query Match 100.0%; Score 22; DB 19; Length 22;  
Best Local Similarity 100.0%; Pred. No. 0.11;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 1 tgttgatcacagaattgata 22  
|||||  
DB 1 tgttgatcacagaattgata 22  
  
RESULT 3  
AAV58896  
ID AAV58896 standard; DNA; 1477 BP.  
XX  
AC AAV58896;  
XX  
DT 20-JAN-1999 (first entry)  
XX  
DE L. fainei nucleotide sequence.  
XX  
KW Infection; pathogenic Leptospira; protective immunity; therapy;  
KM diagnosis; ss.  
XX  
OS Leptospira fainei.  
XX  
PN WO9840099-A1.  
XX  
PD 17-SEP-1998.  
XX  
PE 06-MAR-1998; 98WO-AU00145.  
XX  
PR 07-MAR-1997; 97AU-0005494.  
XX  
PA (AGRI-) AGRIC VICTORIA SERVICES PTY LTD.  
XX (PIGR-) PIG RES & DEV CORP.  
XX  
PI Chappel RJ;  
XX  
DR WPI: 1998-520791/44.  
XX  
PT New isolated pathogenic Leptospira bacterium - useful for, e.g  
PT developing products for conferring protective immunity, and for  
PT prophylactic or therapeutic treatment  
XX  
PS Claim 15; Page 69-70; 94pp; English.  
XX  
CC This sequence represents a Leptospira DNA sequence isolated from the  
CC pathogenic Leptospira (LS) bacterium of the invention. The bacterium  
CC belongs to serogroup Hurstbridge or serovar hurstbridge or the species  
CC L. fainei. The LS bacteria can be used for conferring protective  
CC immunity against pathogenic LS bacteria in humans or animals. The  
CC bacteria can also be used for prophylactic or therapeutic treatment of LS  
CC infections. The DNAs and antibodies may also be used for detection and  
CC diagnosis of past or present LS infection.  
XX  
SQ Sequence 1477 BP; 390 A; 334 C; 439 G; 314 T; 0 other;  
  
Query Match 100.0%; Score 22; DB 19; Length 1477;  
Best Local Similarity 100.0%; Pred. No. 0.16;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 1 tgttgatcacagaattgata 22  
|||||  
DB 154 tgttgatcacagaattgata 175  
  
RESULT 4  
AAA99365  
ID AAA99365 standard; DNA; 1012 BP.  
XX  
AC AAA99365;  
XX  
DT 22-JAN-2001 (first entry)  
XX  
DE Plant PRG1 promoter related gene sequence.

XX Plant promoter; PrAG1; reproductive tissue; transgenic plant; cereal; ds.  
XX Arabidopsis thaliana.  
XX WO200055172-A1.  
XX 21-SEP-2000.  
XX 17-MAR-2000; 2000WO-NZ00031.  
XX 17-MAR-1999; 99NZ-0334715.  
XX (CART-) CARTER HOLT HARVEY LTD.  
XX (TASM-) TASMAN BIOTECHNOLOGY LTD.  
XX (UNMT) UNIV MICHIGAN TECHNOLOGICAL.  
XX Podila GK, Liu J, Karnosky DF;  
XX WPI: 2000-594442/56.  
XX P-PSDB; AAB26796.  
XX Novel plant reproductive tissue promoter, useful to produce plants  
XX which have a diminished reproductive capacity or which are sterile  
XX Claim 15; Page 42-43; 51pp; English.  
XX This invention relates to a novel plant promoter gene. The promoter is  
XX located in plant reproductive tissue, and the invention includes  
XX transgenic plants containing the promoter. The promoter can be used to  
XX produce plants which have a diminished reproductive capacity or which are  
XX sterile. The constructs can also be used to transform agronomically  
XX important plants in which modulation of reproductive capacity  
XX (particularly the timing and abundance of flowering) is desirable,  
XX e.g. cereals, rice, maize, wheat, barley, oats, rye, soybean and canola.  
XX The present sequence represents DNA encoding the plant reproductive  
XX promoter (PrAG1) of the invention.  
XX Sequence 1012 BP; 287 A; 197 C; 217 G; 311 T; 0 other;  
XX  
XX Query Match 85.5%; Score 18.8; DB 21; Length 1012;  
XX Best Local Similarity 90.9%; Pred. No. 5.8;  
XX Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
XX  
XX 1 tgttgatcacaaagtgtgata 22  
XX ||||||| ||||||| |||||||  
XX Db 672 tgttgatcacaaagtgtgata 693  
XX  
XX RESULT 5  
XX AAC37325  
XX ID AAC37325 standard; DNA; 1069 BP.  
XX  
XX AAC37325;  
XX  
XX 17-OCT-2000 (first entry)  
XX  
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 16971.  
XX  
XX Hybridisation assay; genetic mapping; gene expression control;  
XX protein identification; signal transduction pathway;  
XX metabolic pathway; promoter; termination sequence; ss.  
XX Arabidopsis thaliana.  
XX  
XX EP1033405-A2.  
XX  
XX 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 99US-0121825.1  
XX

PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
PR 14-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0135421.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 28-MAY-1999; 99US-0136392.  
PR 01-JUN-1999; 99US-0136782.  
PR 03-JUN-1999; 99US-0137222.  
PR 04-JUN-1999; 99US-0137528.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139859.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140354.  
PR 28-JUN-1999; 99US-0140695.  
PR 29-JUN-1999; 99US-0140823.  
PR 30-JUN-1999; 99US-0140991.  
PR 01-JUL-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145919.  
PR 27-JUL-1999; 99US-0145918.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0146389.  
PR 04-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 06-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.

PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161921.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 85.5%; Score 18.8; DB 21; Length 1069;  
Best Local Similarity 90.9%; Pred. NO. 5.8;  
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Oy 1 ttttgatcacagaattcgata 22  
||||| |||||||  
Db 690 ttttgatcacagaattcgata 711

RESULT 6  
AAAT79696/c  
ID AAAT79696 standard; cDNA: 2084 BP.  
XX  
AC AAAT79696;  
XX  
DT 27-NOV-2000 (first entry)  
DE Eucalyptus grandis cell signalling involved polynucleotide SEQ ID NO:848.  
XX  
KW Eucalyptus grandis; Pinus radiata; Monterey pine; plant; modification;  
KW plant cell signalling; modulation; transgenic plant; pathogen; growth;  
KW environmental change; development; cell proliferation; differentiation;  
KW elongation; survival; disease resistance; nutrient metabolism; ss.  
OS Eucalyptus grandis.  
XX  
XX  
XX WO200042171-A1.  
XX  
XX 20-JUL-2000.  
XX  
XX 11-JAN-2000; 2000WO-US00724.  
XX  
XX 12-JAN-1999; 99US-0228986.  
XX 01-NOV-1999; 99US-0162866.  
XX  
XX (GENE-) GENESIS RES & DEV CORP LTD.  
XX  
XX Strabala TJ, Nieuwehuizen NJ;  
XX

DR WPI: 2000-476052/41.  
XX Isolated polynucleotide encoding a polypeptide involved in cell  
PT signaling used for generating transgenic plants with modified responses  
XX to external signals -  
XX  
XX Claim 1: Page 408-409; 527pp; English.  
XX  
CC AA79263 to AA79736 and AAB25100 to AAB25570 represent polynucleotide  
CC and protein sequences isolated from eucalyptus (Eucalyptus grandis) or  
CC pine (Pinus radiata also known as Monterey pine). The protein sequences  
CC are involved in cell signaling. The polynucleotide and protein  
CC sequences can be used to modify the response of plant cells to external  
CC signals e.g. environmental changes or pathogens during the growth and  
CC development of a plant. They can be used to modify cell proliferation,  
CC differentiation, elongation and survival, resistance to disease and  
CC nutrient metabolism. Examples of modifications which can be produced are  
CC altered fruit ripening and senescence of leaves and flowers e.g. to  
CC delay senescence and prolong the life of cut flowers or enhance  
CC senescence of reproductive organs to engineer sterile plants. Other  
CC modifications can be used to delay senescence in selected cell types or  
CC organs providing fruit and vegetables which have a longer shelf life  
CC between harvest and consumption, or to decrease branching frequency in  
CC forest tree species giving long stretches of valuable knot-free clear  
CC wood which can be used in solid timber furniture and veneers.  
XX  
SQ Sequence 2084 BP; 502 A; 434 C; 566 G; 582 T; 0 other;

Query Match 80.9%; Score 17.8; DB 21; Length 2084;  
Best Local Similarity 90.5%; Pred. No. 19;  
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 tcttgatcaccaagattgac 21  
||||| ||||| |||||  
DB 1347 TGTGTGACGACCAAGTTTGAT 1327.

RESULT 7  
AAC18114  
ID AAC18114 standard; cDNA; 447 BP.  
XX  
AC AAC18114;  
XX  
DT 06-OCT-2000 (first entry)  
XX  
DE Human secreted protein 5' EST, SEQ ID NO: 22189;  
XX  
XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
XX gene therapy; chromosome mapping; ss.  
XX  
OS Homo sapiens.  
XX  
PN EP1033401-A2.  
PD 06-SEP-2000.  
XX  
PF 21-FEB-2000; 2000EP-0200610.  
XX  
PR 26-FEB-1999; 990S-0122487.  
XX  
PA (GEST) GENSET.  
XX  
PI Dumas Mline Edwards J, Duclert A, Giordano J;  
XX  
XX WPI: 2000-500381/45.  
DR  
XX  
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -  
XX  
PS Claim 1; SEQ ID 22189; 71pp + CD-ROM; English.  
XX

CC The present sequence is one of a large number of 5' ESTs derived from  
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively  
CC identified within the present sequence. The 5' ESTs were prepared from  
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST  
CC sequences usually correspond mainly to the 3' untranslated region (UTR)  
CC of the mRNA because they are often obtained from oligo-dT primed cDNA  
CC libraries. Such ESTs are not well suited for isolating cDNA sequences  
CC derived from the 5' ends of mRNAs and even in those cases where longer  
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.  
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be  
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used  
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.  
CC They are used to obtain upstream regulatory sequences and to design  
CC expression and secretion vectors.  
XX  
SQ Sequence 447 BP; 187 A; 57 C; 66 G; 128 T; 9 other;

Query Match 78.2%; Score 17.2; DB 21; Length 447;  
Best Local Similarity 86.4%; Pred. No. 32;  
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 tcttgatcaccaagattgata 22  
||||| ||||| | |||  
DB 240 tcttgatcaccaactata 261

RESULT 8  
AAC42365/C  
ID AAC42365 standard; DNA; 1290 BP.  
XX  
AC AAC42365;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 35275.  
XX  
XX Hybridisation assay; genetic mapping; gene expression control;  
XX protein identification; signal transduction pathway;  
XX metabolic pathway; promoter; termination sequence; ss.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 990S-0121825.  
XX  
PR 05-MAR-1999; 990S-0123180.  
XX  
PR 09-MAR-1999; 990S-0123548.  
XX  
PR 23-MAR-1999; 990S-0125788.  
XX  
PR 25-MAR-1999; 990S-0126264.  
XX  
PR 29-MAR-1999; 990S-0126785.  
XX  
PR 01-APR-1999; 990S-0127462.  
XX  
PR 06-APR-1999; 990S-0128234.  
XX  
PR 08-APR-1999; 990S-0128714.  
XX  
PR 16-APR-1999; 990S-0129645.  
XX  
PR 19-APR-1999; 990S-0130077.  
XX  
PR 21-APR-1999; 990S-0130449.  
XX  
PR 23-APR-1999; 990S-0130510.  
XX  
PR 23-APR-1999; 990S-0130891.  
XX  
PR 28-APR-1999; 990S-0131449.  
XX  
PR 30-APR-1999; 990S-0132048.  
XX  
PR 30-APR-1999; 990S-0132407.  
XX  
PR 04-MAY-1999; 990S-0132484.  
XX  
PR 05-MAY-1999; 990S-0132485.  
XX  
PR 06-MAY-1999; 990S-0132486.  
XX  
PR 07-MAY-1999; 990S-0132487.  
XX  
PR 11-MAY-1999; 990S-0132863.  
XX  
PR 14-MAY-1999; 990S-0134256.  
XX  
PR 14-MAY-1999; 990S-0134218.

PR 14-MAY-1999; 99US-0134219.  
 PR 14-MAY-1999; 99US-0134221.  
 PR 14-MAY-1999; 99US-0134370.  
 PR 18-MAY-1999; 99US-0134768.  
 PR 19-MAY-1999; 99US-0134941.  
 PR 20-MAY-1999; 99US-0135124.  
 PR 21-MAY-1999; 99US-0135529.  
 PR 24-MAY-1999; 99US-0135629.  
 PR 25-MAY-1999; 99US-0136021.  
 PR 27-MAY-1999; 99US-0136392.  
 PR 28-MAY-1999; 99US-0136782.  
 PR 01-JUN-1999; 99US-0137222.  
 PR 03-JUN-1999; 99US-0137528.  
 PR 04-JUN-1999; 99US-0137502.  
 PR 07-JUN-1999; 99US-0137724.  
 PR 08-JUN-1999; 99US-0138094.  
 PR 10-JUN-1999; 99US-0138540.  
 PR 10-JUN-1999; 99US-0138847.  
 PR 14-JUN-1999; 99US-0139119.  
 PR 16-JUN-1999; 99US-0139452.  
 PR 16-JUN-1999; 99US-0139453.  
 PR 17-JUN-1999; 99US-0139492.  
 PR 18-JUN-1999; 99US-0139454.  
 PR 18-JUN-1999; 99US-0139455.  
 PR 18-JUN-1999; 99US-0139456.  
 PR 18-JUN-1999; 99US-0139457.  
 PR 18-JUN-1999; 99US-0139458.  
 PR 18-JUN-1999; 99US-0139459.  
 PR 18-JUN-1999; 99US-0139460.  
 PR 18-JUN-1999; 99US-0139461.  
 PR 18-JUN-1999; 99US-0139462.  
 PR 18-JUN-1999; 99US-0139463.  
 PR 18-JUN-1999; 99US-0139750.  
 PR 18-JUN-1999; 99US-0139763.  
 PR 21-JUN-1999; 99US-0139817.  
 PR 22-JUN-1999; 99US-0139899.  
 PR 23-JUN-1999; 99US-0140353.  
 PR 23-JUN-1999; 99US-0140354.  
 PR 24-JUN-1999; 99US-0140685.  
 PR 28-JUN-1999; 99US-0140828.  
 PR 29-JUN-1999; 99US-0140991.  
 PR 30-JUN-1999; 99US-0141287.  
 PR 01-JUL-1999; 99US-0141642.  
 PR 01-JUL-1999; 99US-0142154.  
 PR 02-JUL-1999; 99US-0142055.  
 PR 06-JUL-1999; 99US-0142390.  
 PR 08-JUL-1999; 99US-0142803.  
 PR 09-JUL-1999; 99US-0142920.  
 PR 12-JUL-1999; 99US-0142977.  
 PR 13-JUL-1999; 99US-0143542.  
 PR 14-JUL-1999; 99US-0143624.  
 PR 15-JUL-1999; 99US-0144005.  
 PR 16-JUL-1999; 99US-0144085.  
 PR 16-JUL-1999; 99US-0144086.  
 PR 19-JUL-1999; 99US-0144325.  
 PR 19-JUL-1999; 99US-0144331.  
 PR 19-JUL-1999; 99US-0144332.  
 PR 19-JUL-1999; 99US-0144333.  
 PR 19-JUL-1999; 99US-0144334.  
 PR 19-JUL-1999; 99US-0144335.  
 PR 20-JUL-1999; 99US-0144352.  
 PR 20-JUL-1999; 99US-0144632.  
 PR 20-JUL-1999; 99US-0144884.  
 PR 21-JUL-1999; 99US-0144814.  
 PR 21-JUL-1999; 99US-0145086.  
 PR 21-JUL-1999; 99US-0145088.  
 PR 22-JUL-1999; 99US-0145085.  
 PR 22-JUL-1999; 99US-0145087.  
 PR 22-JUL-1999; 99US-0145089.  
 PR 22-JUL-1999; 99US-0145192.  
 PR 23-JUL-1999; 99US-0145145.  
 PR 23-JUL-1999; 99US-0145218.  
 PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.  
 PR 27-JUL-1999; 99US-0145913.  
 PR 27-JUL-1999; 99US-0145918.  
 PR 27-JUL-1999; 99US-0145919.  
 PR 28-JUL-1999; 99US-0145951.  
 PR 02-AUG-1999; 99US-0146386.  
 PR 02-AUG-1999; 99US-0146388.  
 PR 03-AUG-1999; 99US-0146389.  
 PR 04-AUG-1999; 99US-0147038.  
 PR 04-AUG-1999; 99US-0147204.  
 PR 04-AUG-1999; 99US-0147302.  
 PR 05-AUG-1999; 99US-0147192.  
 PR 05-AUG-1999; 99US-0147260.  
 PR 06-AUG-1999; 99US-0147303.  
 PR 06-AUG-1999; 99US-0147416.  
 PR 09-AUG-1999; 99US-0147493.  
 PR 09-AUG-1999; 99US-0147935.  
 PR 10-AUG-1999; 99US-0148171.  
 PR 11-AUG-1999; 99US-0148319.  
 PR 12-AUG-1999; 99US-0148341.  
 PR 13-AUG-1999; 99US-0148565.  
 PR 13-AUG-1999; 99US-0148684.  
 PR 16-AUG-1999; 99US-0149368.  
 PR 17-AUG-1999; 99US-0149175.  
 PR 18-AUG-1999; 99US-0149426.  
 PR 20-AUG-1999; 99US-0149722.  
 PR 20-AUG-1999; 99US-0149723.  
 PR 20-AUG-1999; 99US-0149929.  
 PR 23-AUG-1999; 99US-0149902.  
 PR 23-AUG-1999; 99US-0149930.  
 PR 25-AUG-1999; 99US-0150566.  
 PR 26-AUG-1999; 99US-0150884.  
 PR 27-AUG-1999; 99US-0151065.  
 PR 27-AUG-1999; 99US-0151066.  
 PR 27-AUG-1999; 99US-0151080.  
 PR 30-AUG-1999; 99US-0151303.  
 PR 31-AUG-1999; 99US-0151438.  
 PR 01-SEP-1999; 99US-0151930.  
 PR 07-SEP-1999; 99US-0152363.  
 PR 10-SEP-1999; 99US-0153070.  
 PR 13-SEP-1999; 99US-0153758.  
 PR 15-SEP-1999; 99US-0154018.  
 PR 16-SEP-1999; 99US-0154039.  
 PR 20-SEP-1999; 99US-0154779.  
 PR 23-SEP-1999; 99US-0155139.  
 PR 23-SEP-1999; 99US-0155486.  
 PR 24-SEP-1999; 99US-0155559.  
 PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156586.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158359.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 13-OCT-1999; 99US-0159295.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.

```
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match      78.2%; Score 17.2; DB 21; Length 1290;
Best Local Similarity 86.4%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1 tcttgatcaccaagattgata 22
    ||||| 1 |||||1111111111
Db 642 TGTGTGACACACAAAGATTGTGTA 621

RESULT 9
AAC46258/c
ID AAC46258 standard; DNA; 1413 BP.
XX
AC AAC46258;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 49483.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 16-APR-1999; 99US-0128714.
PR 19-APR-1999; 99US-0129845.
PR 21-APR-1999; 99US-0130077.
PR 23-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0144816.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
```

PR 02-AUG-1999; 99US-0146389.  
 PR 03-AUG-1999; 99US-0147038.  
 PR 04-AUG-1999; 99US-0147204.  
 PR 04-AUG-1999; 99US-0147302.  
 PR 05-AUG-1999; 99US-0147192.  
 PR 05-AUG-1999; 99US-0147260.  
 PR 06-AUG-1999; 99US-0147303.  
 PR 06-AUG-1999; 99US-0147416.  
 PR 09-AUG-1999; 99US-0147493.  
 PR 09-AUG-1999; 99US-0147935.  
 PR 10-AUG-1999; 99US-0148171.  
 PR 11-AUG-1999; 99US-0148319.  
 PR 12-AUG-1999; 99US-0148341.  
 PR 13-AUG-1999; 99US-0148565.  
 PR 13-AUG-1999; 99US-0148684.  
 PR 16-AUG-1999; 99US-0149368.  
 PR 17-AUG-1999; 99US-0149175.  
 PR 18-AUG-1999; 99US-0149426.  
 PR 20-AUG-1999; 99US-0149722.  
 PR 20-AUG-1999; 99US-0149723.  
 PR 20-AUG-1999; 99US-0149929.  
 PR 23-AUG-1999; 99US-0149902.  
 PR 23-AUG-1999; 99US-0149930.  
 PR 25-AUG-1999; 99US-0150566.  
 PR 26-AUG-1999; 99US-0150884.  
 PR 27-AUG-1999; 99US-0151065.  
 PR 27-AUG-1999; 99US-0151066.  
 PR 27-AUG-1999; 99US-0151080.  
 PR 30-AUG-1999; 99US-0151303.  
 PR 31-AUG-1999; 99US-0151438.  
 PR 01-SEP-1999; 99US-0151930.  
 PR 07-SEP-1999; 99US-0152363.  
 PR 10-SEP-1999; 99US-0153070.  
 PR 13-SEP-1999; 99US-0153758.  
 PR 15-SEP-1999; 99US-0154018.  
 PR 16-SEP-1999; 99US-0154039.  
 PR 20-SEP-1999; 99US-0154779.  
 PR 22-SEP-1999; 99US-0155139.  
 PR 23-SEP-1999; 99US-0155486.  
 PR 24-SEP-1999; 99US-0155659.  
 PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 04-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 13-OCT-1999; 99US-0159295.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 25-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.

PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

Query Match 78.28; Score 17.2; DB 21; Length 1413;  
 Best Local Similarity 86.4%; Pred. No. 36;  
 Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ttttggatcacagaatttgata 22  
 Db 576 TGTGTGACACAGATTGTGA 555

RESULT 10  
 AAS22799  
 ID AAS22799 standard; cDNA; 2038 BP.  
 XX  
 AC AAS22799;  
 XX  
 DT 24-OCT-2001 (first entry)  
 XX  
 DE Human cDNA encoding a novel human protein #365.  
 XX  
 KW Human: novel protein; ss; Antianaemic; osteopathic; antiinflammatory;  
 KW Immunomodulatory; cytosolic; neuroprotective; vlnetary; nocrotropic;  
 KW anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral;  
 KW antibacterial; antiallergic; dermatological; haemostatic; antisthmatic;  
 KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;  
 KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;  
 KW tissue regeneration; immune disorder.  
 KW  
 XX  
 OS Homo sapiens.  
 OS  
 PN WO200155437-A2.  
 XX  
 PD 02-AUG-2001.  
 XX  
 PF 25-JAN-2001; 2001WO-US02623.  
 XX  
 PR 25-JAN-2000; 2000US-0491404.  
 XX  
 PA (HYSE-) HYSEO INC.  
 XX  
 PI Tang YT, Liu C, Drmanac RT;  
 XX  
 DR WPI; 2001-451939/48.  
 DR P-PSDB; AAU14494.  
 XX  
 PT Isolated polypeptides useful for treating anti-inflammatory diseases,  
 PT nervous system disorders, and for regenerating bone and cartilage -  
 XX  
 Claim 1; Page 733-734; 894pp; English.  
 XX  
 CC The invention relates to polynucleotides encoding novel human  
 CC proteins or their active domains. The polypeptides, polynucleotides and  
 CC antibodies raised against the polypeptides are used in a method of  
 CC treatment of a mammal and prevention of disorders caused by the aberrant  
 CC protein expression or activity. The polypeptides can be used as  
 CC molecular weight markers, food supplements, and in antibody production.  
 CC The polypeptides are used to identify compounds which bind to the  
 CC polypeptides. Polynucleotides of the invention are used as probes and  
 CC primers, for sequencing, for chromosome or gene mapping. In the  
 CC production of recombinant proteins, and in generating anti-sense DNA or  
 CC RNA and in gene therapy. Polypeptides of the invention can be used to  
 CC target drugs to a tumour, in assays to determine biological activity, to  
 CC raise antibodies/elicite an immune response, to determine quantitative  
 CC protein levels, as tissue markers, and to isolate receptors or ligands.  
 CC Polypeptides of the invention may also be useful in treating platelet  
 CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,  
 CC ligament and/or nerve tissue, wound healing, treating burns, promoting  
 CC the proliferation, differentiation and survival of stem cells, as a  
 CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,  
 CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral



CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or  
 CC fungal infection or from autoimmunity, cancer, allergy, asthma,  
 CC graft-versus-host disease, eczema, hemophilia, thrombosis,  
 CC anti-inflammatory diseases, nervous system disorders, and infection.  
 CC The present sequence encodes a protein of the invention.

XX Sequence 2038 BP; 658 A; 400 C; 503 G; 477 T; 0 other;

Query Match 78.2%; Score 17.2; DB 22; Length 2038;

Best Local Similarity 86.4%; Pred. No. 38;

Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 ttttgatcacagaattgata 22  
 ||||| ||||| ||||| |||||

Db 1929 ttttgagacagaattttaa 1950

# RESULT 11

AAH15352/c  
 ID AAH15352 standard; cDNA; 1764 BP.

AC AAH15352;

DT 26-JUN-2001 (first entry)

DE Human cDNA sequence SEQ ID NO:13526.

XX Human; primer: detection; diagnosis; antisense therapy; gene therapy; ss.

KW Homo sapiens.

OS EPI074617-A2.

PN 07-FEB-2001.

PD 28-JUL-2000; 2000EP-0116126.

PF 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.

PR 11-JAN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

PA (HELI-) HELIX RES INST.

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

DR WPI; 2001-318749/34.

PS Claim 8; SEQ ID 13526; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602  
 CC full-length cDNAs defined in the specification. Where a primer set  
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
 CC to the complementary strand of a polynucleotide which comprises one of  
 CC the 5602 nucleotide sequences defined in the specification, where the  
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in  
 CC the specification. The primer sets can be used in antisense therapy and  
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
 CC particularly full-length cDNAs. The primers are also useful for the

CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB22446 to  
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 CC of the present invention.

XX Sequence 1764 BP; 511 A; 347 C; 392 G; 514 T; 0 other;

Query Match 76.4%; Score 16.8; DB 22; Length 1764;

Best Local Similarity 90.0%; Pred. No. 58;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ttttgatcacagaattgata 20  
 ||||| ||||| ||||| |||||

Db 783 ttttgatcacagaatttga 764

# RESULT 12

AAT90543/c  
 ID AAT90543 standard; DNA; 2582 BP.

AC AAT90543;

DT 13-FEB-1998 (first entry)

DE Maize male fertility gene Z33 Zm41-A.

XX Zm41-A; male fertility; hybrid seed; transgenic plant; antisense;

KW ribozyme; male sterile; maize; Ms41-A; ds.

XX

OS

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

## Location/Qualifiers

Key 823..915

FT exon /\*tag= a

FT intron 916..935

FT exon /\*tag= b

FT exon 936..1090

FT intron /\*tag= c

FT intron 1091..1167

FT exon /\*tag= d

FT exon 1168..1332

FT exon /\*tag= e

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Claim 3; Fig 15; 85pp; English.

This DNA sequence comprises the maize Z33 Zm41-A gene. This is an

orthologue of Arabidopsis Ms41-A (see AAT90522), a gene that confers

male fertility. It was obtained following a database search for

sequences that showed homology to Ms41-A DNA. Isolation of a

partial cDNA clone, and use of this clone to isolate Zm41-A genes

Z31 (AAT90542), Z33 (AAT90543) and Z35 (AAT90544) from genomic lambda

libraries. The genes show a high level of conservation. Z35 may

CC be derived from 231 via genetic rearrangements, deletions and/or  
CC insertions. 233 has subsequent deletions from 235 and is  
CC truncated, having only exons 3, 5 and 6. The MS41-A and Zm41-A  
CC genes, antisense or ribozyme sequences can be used to produce  
CC transgenic plants with controlled male fertility. Male sterile  
CC plants are useful for hybrid seed production, particularly in  
CC Brassicaceae, lettuce, spinach and onions.  
CC  
XX  
SQ Sequence 2582 BP; 670 A; 512 C; 537 G; 863 T; 0 other;  
  
Query Match 76.4%; Score 16.8; DB 18; Length 2582;  
Best Local Similarity 90.0%; Pred. No. 60;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
OY 1 tggtgacacacagatttga 20  
Db 268 TGTGTGATTCACAGATCTGA 249  
||||| ||||||| |||  
  
RESULT 13  
AAC56059  
ID AAC56059 standard; DNA: 2444 BP.  
AC AAC56059;  
XX  
XX  
DT 25-JAN-2001 (first entry)  
XX  
XX  
DE Eucalyptus grandis transcription factor DNA sequence #190.  
XX  
XX  
KW Plant; transcription factor; gene expression; eucalyptus; pine; acacia;  
KW poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;  
KW basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;  
KW homeodomain zipper; LIM domain; AP2; EREBs; zinc finger domain;  
KW type 2 Cys2His2; CCAAT box element; MYB; ss.  
XX  
XX  
OS Eucalyptus grandis.  
XX  
XX  
PN WO200053724-A2.  
XX  
XX  
PD 14-SEP-2000.  
XX  
XX  
PF 09-MAR-2000; 2000MO-US06112.  
XX  
XX  
PR 11-MAR-1999; 99US-0266513.  
PR 18-AUG-1999; 99US-0149485.  
XX  
XX  
PA (GENE-) GENESIS RES & DEV CORP LTD.  
PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.  
XX  
XX  
PI Wood M, McGrath A, Shenk MA, Glenn M;  
XX  
XX  
DR WPI: 2000-579369/54.  
XX  
XX  
PT New isolated polynucleotide encoding a plant transcription factor for  
PT producing a plant e.g. a woody plant, preferably eucalyptus or pine,  
PT having modified gene expression or modified activity of a polypeptide  
PT  
XX  
XX  
PS Claim 1; Pages 94-95; 747pp; English.  
XX  
XX  
CC The present invention relates to novel plant transcription factors from  
CC Eucalyptus grandis or Pinus radiata. The present sequence is the coding  
CC sequence for one such transcription factor. The transcription factor may  
CC be used to produce a plant having modified gene expression such as a  
CC woody plant e.g. a eucalyptus, pine, acacia, poplar, sweetgum, teak, or  
CC mahogany species or to modify the activity of a polypeptide in a plant.  
CC The transcription factors of the present invention are members from the  
CC following families of regulatory proteins: bZIP, bZIP family of G-box  
CC binding factors, basic helix-loop-helix zipper,  
CC homeotic/homeodomain/homeobox/MADS, homeodomain zipper, LIM domain, AP2  
CC and EREBs, zinc finger domains of type 2 Cys2His2, CCAAT box elements

CC and MYB.  
XX  
SQ Sequence 2444 BP; 643 A; 554 C; 640 G; 607 T; 0 other;

Query Match 74.5%; Score 16.4; DB 21; Length 2444;  
Best Local Similarity 94.4%; Pred. No. 94;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 tggtgacacacagattt 18  
Db 1451 tgatgacacacagattt 1468  
||||| ||||||| |||

RESULT 14  
AAH33702/C  
ID AAH33702 standard; cDNA; 335 BP.  
AC AAH33702;  
XX  
XX  
DT 03-SEP-2001 (first entry)  
XX  
XX  
DE Human colon cancer antigen encoding cDNA SEQ ID NO:758.  
XX  
XX  
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;  
KW colorectal carcinoma; ss.  
XX  
XX  
OS Homo sapiens.  
XX  
XX  
PN WO200122920-A2.  
XX  
XX  
PD 05-APR-2001.  
XX  
XX  
PF 28-SEP-2000; 2000MO-US26524.  
XX  
XX  
PR 29-SEP-1999; 99US-0157137.  
PR 03-NOV-1999; 99US-0163280.  
XX  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX  
PI Ruben SM, Barash SC, Birse CE, Rosen CA;  
XX  
XX  
DR WPI: 2001-235357/24.  
DR P-PSDB; AAG74271.  
XX  
XX  
PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,  
PT useful for preventing, diagnosing and/or treating colorectal cancers -  
PT  
XX  
XX  
PS Claim 1; Page 2738; 9803pp; English.  
XX  
XX  
CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon  
CC cancer-associated nucleic acid molecules (N) and proteins (P), where  
CC the proteins are collectively known as colon cancer antigens. The colon  
CC cancer antigens have cytostatic activity and can be used in gene  
CC therapy and vaccine production. N and P may be used in the prevention,  
CC diagnosis and treatment of diseases associated with inappropriate P  
CC expression. For example, N and P may be used to treat disorders  
CC associated with decreased expression by rectifying mutations or deletions  
CC in a patient's genome that affect the activity of P by expressing  
CC inactive proteins or to supplement the patients own production of P.  
CC Additionally, N may be used to produce the colon cancer-associated Ps  
CC by inserting the nucleic acids into a host cell and culturing the cell  
CC to express the proteins. N and P can be used in the prevention, diagnosis  
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204  
CC and AAB77789 represent sequences used in the exemplification of the  
CC present invention.  
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were  
CC missing at time of publication, meaning no sequences are present for  
CC SEQ ID NO:1027 to 1052, 7921 and 7922.  
XX  
XX  
SQ Sequence 335 BP; 95 A; 71 C; 60 G; 105 T; 4 other;

```

Query Match          73.6%; Score 16.2; DB 22; Length 335;
Best Local Similarity 85.7%; Pred. No. 96;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY      1 ttttgatcacagaatttcg 21
        ||||| | ||||| ||||| ||
Db       122 TGTGTGACGACAAAGATTTCAT 102

RESULT 15
AAV84598/1
ID      AAV84598 standard; DNA; 951 BP.
XX
XX      AAV84598;
AC
XX
XX      01-MAR-1999 (first entry)
DT
XX
DE      Human secreted protein gene 188 clone HHP5F70.
XX
XX      Human; secreted protein; fusion protein; gene therapy; protein therapy;
XX      diagnosis; cancer; tumour; neurodegenerative disorder; leukaemia;
XX      developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
XX      immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
XX      inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
XX      cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
XX      osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
XX      endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
OS      Homo sapiens.
XX
XX      W09854963-A2.
XX
XX      10-DEC-1998.
XX
XX      04-JUN-1998; 98MO-US11422.
XX
XX      18-DEC-1997; 97US-0070923.
XX      06-JUN-1997; 97US-0048877.
XX      06-JUN-1997; 97US-0048881.
XX      06-JUN-1997; 97US-0048884.
XX      06-JUN-1997; 97US-0048893.
XX      06-JUN-1997; 97US-0048896.
XX      06-JUN-1997; 97US-0048899.
XX      06-JUN-1997; 97US-0048915.
XX      06-JUN-1997; 97US-0048949.
XX      06-JUN-1997; 97US-0048964.
XX      06-JUN-1997; 97US-0048972.
XX      06-JUN-1997; 97US-0049020.
XX      05-SEP-1997; 97US-0049375.
XX      05-SEP-1997; 97US-0057628.
XX      05-SEP-1997; 97US-0057635.
XX      05-SEP-1997; 97US-0057644.
XX      05-SEP-1997; 97US-0057647.
XX      05-SEP-1997; 97US-0057660.
XX      05-SEP-1997; 97US-0057661.
XX      05-SEP-1997; 97US-0057667.
XX      05-SEP-1997; 97US-0057761.
XX      05-SEP-1997; 97US-0057764.
XX      05-SEP-1997; 97US-0057770.
XX      05-SEP-1997; 97US-0057775.
XX      05-SEP-1997; 97US-0057778.
XX      06-JUN-1997; 97US-0048875.
XX      06-JUN-1997; 97US-0048878.
XX      06-JUN-1997; 97US-0048882.
XX      06-JUN-1997; 97US-0048885.
XX      06-JUN-1997; 97US-0048894.
XX      06-JUN-1997; 97US-0048897.
XX      06-JUN-1997; 97US-0048900.
XX      06-JUN-1997; 97US-0048916.
XX      06-JUN-1997; 97US-0048962.
XX      06-JUN-1997; 97US-0048970.
XX      06-JUN-1997; 97US-0048974.
XX      06-JUN-1997; 97US-0049373.

```

```

PR      05-SEP-1997; 97US-0057584.
PR      05-SEP-1997; 97US-0057629.
PR      05-SEP-1997; 97US-0057642.
PR      05-SEP-1997; 97US-0057645.
PR      05-SEP-1997; 97US-0057648.
PR      05-SEP-1997; 97US-0057651.
PR      05-SEP-1997; 97US-0057662.
PR      05-SEP-1997; 97US-0057668.
PR      05-SEP-1997; 97US-0057762.
PR      05-SEP-1997; 97US-0057765.
PR      05-SEP-1997; 97US-0057771.
PR      05-SEP-1997; 97US-0057776.
PR      06-JUN-1997; 97US-0048876.
PR      06-JUN-1997; 97US-0048880.
PR      06-JUN-1997; 97US-0048883.
PR      06-JUN-1997; 97US-0048892.
PR      06-JUN-1997; 97US-0048895.
PR      06-JUN-1997; 97US-0048898.
PR      06-JUN-1997; 97US-0048901.
PR      06-JUN-1997; 97US-0048917.
PR      06-JUN-1997; 97US-0048963.
PR      06-JUN-1997; 97US-0048971.
PR      06-JUN-1997; 97US-0049019.
PR      06-JUN-1997; 97US-0049374.
PR      05-SEP-1997; 97US-0057627.
PR      05-SEP-1997; 97US-0057634.
PR      05-SEP-1997; 97US-0057643.
PR      05-SEP-1997; 97US-0057646.
PR      05-SEP-1997; 97US-0057649.
PR      05-SEP-1997; 97US-0057654.
PR      05-SEP-1997; 97US-0057666.
PR      05-SEP-1997; 97US-0057760.
PR      05-SEP-1997; 97US-0057763.
PR      05-SEP-1997; 97US-0057769.
PR      05-SEP-1997; 97US-0057774.
PR      05-SEP-1997; 97US-0057777.
XX
XX      (HUMA-) HUMAN GENOME SCI INC.
XX
XX      Brewer LA, Carter KC, Dillon PJ, Edner R, Endress GA;
XX      Fan P, Feng P, Ferlie AM, Fischer CU, Florence D;
XX      Florene K, Greene JM, Hu J, Kyaw H, Lafleur DW;
XX      Li Y, Moore PA, Ni J, Olsen HS, Rosen CA, Ruben SM;
XX      Shi Y, Soppet DR, Wei Y, Young P, Yu G, Zeng Z;
XX      MPI: 1999-059865/05.
XX      P-PSDB: AAW88721.
XX
XX      New isolated human genes and the secreted polypeptides they encode -
XX      useful for diagnosis and treatment of e.g. cancers, neurological
XX      disorders, immune diseases, inflammation or blood disorders
XX
XX      Claim 4; Page 450-451; 772pp; English.
XX
XX      The invention relates to nucleic acid sequences (AAV84411 to AAV84633)
XX      encoding human secreted proteins (AAW88534 to AAW88756). The secreted
XX      protein gene sequences are deposited with the ATCC under deposit numbers
XX      ATCC 97979, 97974, 97975, 97976, 97977, 209007, 209008, 209009, 209010,
XX      209011, 209080, 209081, 209082, 209083, 209084, 209085, 209511. Host
XX      cells comprising recombinant vectors containing the nucleic acid
XX      sequences are used for the recombinant production of the secreted
XX      proteins. The polynucleotide and amino acid sequences are useful for are
XX      useful for preventing, treating or ameliorating medical conditions e.g.
XX      by protein or gene therapy. Pathological conditions can be also
XX      diagnosed by determining the amount of the new polypeptides in a sample
XX      or by determining the presence of mutations in the new polynucleotides.
XX      Specific uses are described for each of the polynucleotides, based on
XX      which tissues they are most highly expressed in, and include developing
XX      products for the diagnosis or treatment of cancer, neurodegenerative
XX      disorders, developmental abnormalities and foetal deficiencies, blood
XX      disorders, tumours, leukemias, diseases of the immune system, autoimmune
XX      diseases, hepatic and renal disease, lymphomas, inflammation, allergies,
XX      ischemic shock, Alzheimer's and cognitive disorders, schizophrenia,

```

CC restenosis, prostate diseases, obesity, disorders involving osteoclasts  
 CC such as osteoporosis, arthritis or malignancies, diseases of testes,  
 CC lung or thymus, digestive/endocrine disorders, infections and AIDS. The  
 CC polypeptides are also useful for identifying their binding partners.  
 CC The present sequence represents a gene encoding a human secreted protein  
 CC (see descriptor line for gene number and clone identification).  
 XX

Sequence 951 BP; 296 A; 134 C; 192 G; 329 T; 0 other;

Query Match 73.68; Score 16.2; DB 20; Length 951;

Best Local Similarly 85.7%; Pred. No. 1.1e+02;

Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 tgttgatcacagaattgat 21

||||| || ||||| |||||

Db 524 TGTGTGTTCTCAGAGCTCGAT 504

Search completed: December 15, 2001, 03:29:43  
 Job time: 7883 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 03:22:44 : Search time 172.39 Seconds  
(without alignments)  
28.903 Million cell updates/sec

Title: US-09-380-826A-2

Perfect score: 22

Sequence: 1 tctgtgacacagaattgata 22

Scoring table: IDENTITY\_NUC

Gapop 10.0, Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16.8	76.4	2582	4	US-09-102-528-31
2	16.2	73.6	2000	1	US-08-014-326-1
3	16.2	73.6	2001	3	US-08-417-089-1
4	16.2	73.6	2001	3	US-08-695-651-1
5	16.2	73.6	2001	4	US-08-930-285-1
6	16.2	73.6	2001	4	US-08-695-421-1
7	16.2	73.6	4345	2	US-08-244-537-1
8	16.2	73.6	7470	3	US-08-417-089-5
9	16.2	73.6	7470	3	US-08-695-651-5
10	16.2	73.6	7470	4	US-08-930-285-5
11	16.2	73.6	7470	4	US-08-695-421-5
12	15.6	70.9	296	4	US-08-602-145-14
13	15.6	70.9	1095	4	US-08-928-383B-3
14	15.6	70.9	1096	4	US-08-684-862-8
15	15.6	70.9	1584	4	US-08-928-383B-1
16	15.6	70.9	1989	4	US-08-792-055-1
17	15.6	70.9	2434	4	US-09-272-496-1
18	15.6	70.9	2770	4	US-09-008-697A-13
19	15.6	70.9	7739	4	US-09-195-966-1
20	15.6	70.9	7745	4	US-08-299-675-1
21	15.6	70.9	7745	2	US-08-485-241-1
22	15.2	69.1	1016	1	US-08-874-162-1
23	15.2	69.1	1016	1	US-08-399-986B-3
24	15.2	69.1	1016	1	US-08-493-754A-3
25	15.2	69.1	2182	1	US-08-399-986B-1
26	15.2	69.1	2182	1	US-08-493-754A-1
27	15.2	69.1	2598	3	US-08-745-892-20

28	15.2	69.1	5420	6	5256642-3	Patent No. 5256642
29	15.2	69.1	5420	6	5472939-3	Patent No. 5472939
30	15.2	69.1	5872	4	US-09-102-528-32	Sequence 32, Appl
31	15.2	69.1	6951	6	5256642-1	Patent No. 5256642
32	15.2	69.1	6951	6	5472939-1	Patent No. 5472939
33	14.8	67.3	337	4	US-08-991-789A-13	Sequence 13, Appl
34	14.8	67.3	1855	3	US-08-961-083-71	Sequence 71, Appl
35	14.8	67.3	3159	1	US-08-119-361-4	Sequence 4, Appl
36	14.8	67.3	3159	3	US-08-336-308A-3	Sequence 3, Appl
37	14.8	67.3	3159	3	US-08-822-324-3	Sequence 3, Appl
38	14.8	67.3	3159	4	US-09-490-931-3	Sequence 3, Appl
39	14.8	67.3	7266	3	US-08-336-308A-9	Sequence 9, Appl
40	14.8	67.3	7266	3	US-08-822-324-5	Sequence 5, Appl
41	14.8	67.3	7266	4	US-09-490-931-9	Sequence 9, Appl
42	14.8	67.3	8640	1	US-08-570-311-28	Sequence 28, Appl
43	14.8	67.3	9919	3	US-08-880-179-1	Sequence 1, Appl
44	14.6	66.4	293	3	US-08-866-340-13	Sequence 13, Appl
45	14.6	66.4	293	4	US-09-103-875-17	Sequence 17, Appl

## ALIGNMENTS

```
RESULT 1
US-09-102-528-31/c
: Sequence 31, Application US/09102528
: Patent No. 6207883
: GENERAL INFORMATION:
: APPLICANT:
: TITLE OF INVENTION: DNA Sequences
: NUMBER OF SEQUENCES: 32
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC Compatible
: SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/102,528
: FILING DATE:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: WO PCT/GB96/03191
: FILING DATE:
: INFORMATION FOR SEQ ID NO: 31:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2582 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: Single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
US-09-102-528-31

Query Match      76.4%; Score 16.8; DB 4; Length 2582;
Best Local Similarity 90.0%; Pred. No. 14;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY      1 tctgtgacacagaattga 20
Db      268 TGTGATCACAAGATCTGA 249

RESULT 2
US-08-014-326-1
: Sequence 1, Application US/08014326
: Patent No. 548544
: GENERAL INFORMATION:
: APPLICANT: Genenbach, Burle G.
: APPLICANT: Somers, David A.
: APPLICANT: Wyse, Donald L.
: APPLICANT: Gronwald, John W.
: APPLICANT: Egli, Margaret A.
: APPLICANT: Lutz, Shiela M.
: TITLE OF INVENTION: Method and An Acetyl CoA Carboxylase
```

TITLE OF INVENTION: Alteration in Oil Content of Plants  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Merchant & Gould  
STREET: 3100 No. 5498544west Center  
CITY: Minneapolis  
STATE: MN  
COUNTRY: USA  
ZIP: 55402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/014,326  
FILING DATE: 05-FEB-1993  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/917,462  
FILING DATE: 21-JUL-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/538,674  
FILING DATE: 18-JUN-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/269,584  
FILING DATE: 10-NOV-1988  
ATTORNEY/AGENT INFORMATION:  
NAME: Kowalchuk, Katherine M.  
REGISTRATION NUMBER: 36,848  
REFERENCE/DOCKET NUMBER: 600.258-US01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 612-332-5300  
TELEFAX: 612-332-9081  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2000 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
IMMEDIATE SOURCE:  
CLONE: 2 kb fragment of lambda clone #15-14  
US-08-014-326-1

Query Match 73.6%; Score 16.2; DB 1; Length 2000;  
Best Local Similarity 85.7%; Pred. No. 27;  
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 2 gtggatcacagaattgata 22  
||||| ||||| ||| ||  
Db 42 GTTGATGACACAGATTGTTA 62

RESULT 3  
US-08-417-089-1  
Sequence 1, Application US/08417089  
Patent No. 6069298  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: METHODS AND AN ACETYL COA CARBOXYLASE GENE  
TITLE OF INVENTION: FOR CONFERRING HERBICIDE TOLERANCE AND AN ALTERATION IN  
TITLE OF INVENTION: OIL CONTENT OF PLANTS  
NUMBER OF SEQUENCES: 11  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/417,089  
INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:  
LENGTH: 2001 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-417-089-1

Query Match 73.6%; Score 16.2; DB 3; Length 2001;  
Best Local Similarity 85.7%; Pred. No. 27;  
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 2 gtggatcacagaattgata 22  
||||| ||||| ||| ||  
Db 42 GTTGATGACACAGATTGTTA 62

RESULT 4  
US-08-695-651-1  
Sequence 1, Application US/08695651  
Patent No. 6146867  
GENERAL INFORMATION:  
APPLICANT: Gengenbach, B. G.  
APPLICANT: Somers, D. A.  
APPLICANT: Wyse, D. L.  
APPLICANT: Gronwald, J. W.  
APPLICANT: Egli, M. A.  
APPLICANT: Lutz, S. M.  
TITLE OF INVENTION: METHOD FOR PRODUCING PLANT ACETYL COA  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.  
STREET: P.O. Box 2938  
CITY: Minneapolis  
STATE: MN  
COUNTRY: USA  
ZIP: 55402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTED Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/695,651  
FILING DATE: 12-AUG-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/417089  
FILING DATE: 05-APR-1995  
APPLICATION NUMBER: 08/014326  
FILING DATE: 05-FEB-1993  
APPLICATION NUMBER: 07/917462  
FILING DATE: 21-JUL-1992  
APPLICATION NUMBER: 07/538674  
FILING DATE: 18-JUN-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Woessner, Warren D  
REGISTRATION NUMBER: 30,440  
REFERENCE/DOCKET NUMBER: 600.318US3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 612-373-6900  
TELEFAX: 612-339-3061  
TELEX:  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2001 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-08-695-651-1

Query Match 73.6%; Score 16.2; DB 3; Length 2001;  
Best Local Similarity 85.7%; Pred. No. 27;  
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 gttggtcacacaagattgtgata 22  
||||| ||||| ||| ||  
DB 42 GTTGATGACACAGAGTTGTGTA 62

RESULT 5  
US-08-930-285-1  
; Sequence 1, Application US/08930285  
; Patent No. 6222099  
; GENERAL INFORMATION:  
; APPLICANT: Regents of the University of Minnesota, et al.  
; TITLE OF INVENTION: TRANSGENIC PLANTS EXPRESSING ACETYL COA CARBOXYLASE GE  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P. A.  
; STREET: P. O. Box 2938  
; CITY: Minneapolis  
; STATE: MN  
; COUNTRY: USA  
; ZIP: 55402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/930,285  
; FILING DATE: 13-APR-1998  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US96/04625  
; FILING DATE: 04-APR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Embretson, Janet E.  
; REGISTRATION NUMBER: 39,665  
; REFERENCE/DOCKET NUMBER: 600.318US4  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 612-339-0331  
; TELEFAX: 612-339-3061  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2001 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHEICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE:  
; ORIGINAL SOURCE:  
; US-08-930-285-1

Query Match 73.6%; Score 16.2; DB 4; Length 2001;  
Best Local Similarity 85.7%; Pred. No. 27;  
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 gttggtcacacaagattgtgata 22  
||||| ||||| ||| ||  
DB 42 GTTGATGACACAGAGTTGTGTA 62

RESULT 6  
US-08-695-421-1  
; Sequence 1, Application US/08695421  
; Patent No. 6268550  
; GENERAL INFORMATION:

APPLICANT: Gengenbach, B. G.  
APPLICANT: Somers, D. A.  
APPLICANT: Wyse, D. L.  
APPLICANT: Gronwald, J. W.  
APPLICANT: Egli, M. A.  
APPLICANT: Lutz, S. M.  
TITLE OF INVENTION: METHODS AND AN ACETYL COA CARBOXYLASE GENE  
TITLE OF INVENTION: FOR CONFERRING HERBICIDE TOLERANCE AND AN ALTERATION IN OIL  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.  
STREET: P.O. Box 2938  
CITY: Minneapolis  
STATE: MN  
COUNTRY: USA  
ZIP: 55402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/695,421  
FILING DATE: 23-AUG-1996  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/417089  
FILING DATE: 05-APR-1995  
APPLICATION NUMBER: 08/014326  
FILING DATE: 05-FEB-1993  
APPLICATION NUMBER: 07/917462  
FILING DATE: 21-JUL-1992  
APPLICATION NUMBER: 07/538674  
FILING DATE: 18-JUN-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Woessner, Warren D  
REGISTRATION NUMBER: 30,440  
REFERENCE/DOCKET NUMBER: 600.318US2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 612-373-6900  
TELEFAX: 612-339-3061  
TELEX:  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2001 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-695-421-1

Query Match 73.6%; Score 16.2; DB 4; Length 2001;  
Best Local Similarity 85.7%; Pred. No. 27;  
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 gttggtcacacaagattgtgata 22  
||||| ||||| ||| ||  
DB 42 GTTGATGACACAGAGTTGTGTA 62

RESULT 7  
US-08-244-537-1  
; Sequence 1, Application US/08244537  
; Patent No. 5854420  
; GENERAL INFORMATION:  
; APPLICANT: ASHTON, ANTHONY R.  
; APPLICANT: JENKINS, COLIN L.D.  
; APPLICANT: WHITFIELD, PAUL R.  
; TITLE OF INVENTION: MAIZE ACETYL COA CARBOXYLASE ENCODING  
; TITLE OF INVENTION: DNA CLONES  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:

ADDRESSEE: CUSHMAN DAREY & CUSHMAN, L.L.P.  
STREET: 1100 New York Avenue, N.W.  
CITY: Washington  
STATE: D. C.  
COUNTRY: U.S.A.  
ZIP: 20005-3918  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/244,537  
FILING DATE: 18-AUG-1994  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9125330.2  
FILING DATE: 28-NOV-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB92/02205  
FILING DATE: 27-NOV-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: KOKULIS, PAUL N.  
REGISTRATION NUMBER: 16,773  
REFERENCE/DOCKET NUMBER: 203094/SEE 36663/UST  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 861-3000  
TELEFAX: (202) 822-0944  
TELEX: 6714627 CUSH  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4345 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: both  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-244-537-1

Query Match 73.6%; Score 16.2; DB 2; Length 4345;  
Best Local Similarity 85.7%; Pred. No. 30;  
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 gtggatcacagaatttgata 22  
||||| ||||| ||| |||  
Db 902 GTGGATGACACAAGTGTCTTA 922

RESULT 8  
US-08-417-089-5  
Sequence 5, Application US/08417089  
Patent No. 6069298  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: METHODS AND AN ACETYL COA CARBOXYLASE GENE  
TITLE OF INVENTION: FOR CONFERRING HERBICIDE TOLERANCE AND AN ALTERATION IN  
NUMBER OF SEQUENCES: 11  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/417,089  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7470 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-417-089-5

Query Match 73.6%; Score 16.2; DB 3; Length 7470;  
Best Local Similarity 85.7%; Pred. No. 32;  
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 gtggatcacagaatttgata 22  
||||| ||||| ||| |||  
Db 3972 GTTGATGACACAAGTGTCTTA 3992

RESULT 9  
US-08-695-651-5  
Sequence 5, Application US/08695651  
Patent No. 6146867  
GENERAL INFORMATION:  
APPLICANT: Gengenbach, B. G.  
APPLICANT: Somers, D. A.  
APPLICANT: Wyse, D. L.  
APPLICANT: Gronwald, J. W.  
APPLICANT: Egli, M. A.  
APPLICANT: Lutz, S. M.  
TITLE OF INVENTION: METHOD FOR PRODUCING PLANT ACETYL COA  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.  
STREET: P.O. Box 2938  
CITY: Minneapolis  
STATE: MN  
COUNTRY: USA  
ZIP: 55402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/695,651  
FILING DATE: 12-AUG-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/417089  
FILING DATE: 05-APR-1995  
APPLICATION NUMBER: 08/014326  
FILING DATE: 05-FEB-1993  
APPLICATION NUMBER: 07/917462  
FILING DATE: 21-JUL-1992  
APPLICATION NUMBER: 07/538674  
FILING DATE: 18-JUN-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Woessner, Warren D  
REGISTRATION NUMBER: 30,440  
REFERENCE/DOCKET NUMBER: 600.318US3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 612-373-6900  
TELEFAX: 612-339-3061  
TELEX:  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7470 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-08-695-651-5

Query Match 73.6%; Score 16.2; DB 3; Length 7470;  
Best Local Similarity 85.7%; Pred. No. 32;  
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 gtggatcacagaatttgata 22  
||||| ||||| ||| |||



Db 3972 GTTGATGACAGAGTTGTTA 3992

RESULT 10

US-08-930-285-5

Sequence 5, Application US/08930285

Patent No. 6222099

GENERAL INFORMATION:

APPLICANT: Regents of the University of Minnesota, et al.

TITLE OF INVENTION: TRANSGENIC PLANTS EXPRESSING ACETYL COA CARBOXYLASE GE

NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESS:

ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P. A.

STREET: P. O. Box 2938

CITY: Minneapolis

STATE: MN

COUNTRY: USA

ZIP: 55402

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/930,285

FILING DATE: 13-APR-1998

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US96/04625

FILING DATE: 04-APR-1996

ATTORNEY/AGENT INFORMATION:

NAME: Embretson, Janet E.

REGISTRATION NUMBER: 39,665

REFERENCE/DOCKET NUMBER: 600.318054

TELECOMMUNICATION INFORMATION:

TELEPHONE: 612-339-0331

TELEFAX: 612-339-3061

TELEX:

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 7470 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE:

ORIGINAL SOURCE:

US-08-930-285-5

Query Match 73.6%; Score 16.2; DB 4; Length 7470;

Best Local Similarity 85.7%; Pred. No. 32;

Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 gttgatacaagaattgata 22

Db 3972 GTTGATGACAGAGTTGTTA 3992

RESULT 11

US-08-695-421-5

Sequence 5, Application US/08695421

Patent No. 6268550

GENERAL INFORMATION:

APPLICANT: Gengenbach, B. G.

APPLICANT: Somers, D. A.

APPLICANT: Wyse, D. L.

APPLICANT: Gronwald, J. W.

APPLICANT: Egli, M. A.

APPLICANT: Lutz, S. M.

TITLE OF INVENTION: METHODS AND AN ACETYL COA CARBOXYLASE GENE

TITLE OF INVENTION: FOR CONFERRING HERBICIDE TOLERANCE AND AN ALTERATION IN OIL

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P. A.

STREET: P. O. Box 2938

CITY: Minneapolis

STATE: MN

COUNTRY: USA

ZIP: 55402

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/695,421

FILING DATE: 23-AUG-1996

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/417089

FILING DATE: 05-APR-1995

APPLICATION NUMBER: 08/014326

FILING DATE: 05-FEB-1993

APPLICATION NUMBER: 07/917462

FILING DATE: 21-JUL-1992

APPLICATION NUMBER: 07/538674

FILING DATE: 18-JUN-1990

ATTORNEY/AGENT INFORMATION:

NAME: Woessner, Warren D

REGISTRATION NUMBER: 30,440

REFERENCE/DOCKET NUMBER: 600.318052

TELECOMMUNICATION INFORMATION:

TELEPHONE: 612-373-6900

TELEFAX: 612-339-3061

TELEX:

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 7470 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

US-08-695-421-5

Query Match 73.6%; Score 16.2; DB 4; Length 7470;

Best Local Similarity 85.7%; Pred. No. 32;

Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 gttgatacaagaattgata 22

Db 3972 GTTGATGACAGAGTTGTTA 3992

RESULT 12

US-08-602-145-14/C

Sequence 14, Application US/08602145

Patent No. 6025336

GENERAL INFORMATION:

APPLICANT: Goltzy, Kristin L.

APPLICANT: Greenberger, Joel S.

TITLE OF INVENTION: DETERMINING EXPOSURE TO IONIZING RADIATION

TITLE OF INVENTION: AGENT WITH PERSISTENT BIOLOGICAL MARKERS

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: Foley & Lardner

STREET: 3000 K Street, N.W., suite 500

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20007-5109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/602,145  
FILING DATE: 15-FEB-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Bent, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 076333/0112  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 296 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-602-145-14

Query Match 70.9%; Score 15.6; DB 3; Length 296;  
Best Local Similarity 81.8%; Pred. No. 42;  
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ttttgatcacagaatttgata 22  
|||||  
DB 288 TTTTCATCATTAACATTTGATTA 267

RESULT 13  
US-08-928-383B-3/c  
Sequence 3, Application US/p8928383B  
Patent No. 6210921  
GENERAL INFORMATION:  
APPLICANT: Robert W. Flinberg, Jeffrey M. Bergelson,  
APPLICANT: and Marshall S. Horwitz  
TITLE OF INVENTION: CAR, A NO. 6210921el Cocksacklevirus and Adenovirus  
TITLE OF INVENTION: Receptor  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD, LLP  
STREET: 28 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/928,383B  
FILING DATE: 12-SEP-1997  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/026,100  
FILING DATE: 13-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Mandragouras, Amy E.  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: DFN-020  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)742-4214  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1095 base pairs  
TYPE: nucleic acid

STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1095  
US-08-928-383B-3

Query Match 70.9%; Score 15.6; DB 4; Length 1095;  
Best Local Similarity 81.8%; Pred. No. 49;  
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ttttgatcacagaatttgata 22  
|||||  
DB 312 TGATGCATCACAGATTGAGA 291

RESULT 14  
US-08-684-862-8/c  
Sequence 8, Application US/08684862  
Patent No. 5759541  
GENERAL INFORMATION:  
APPLICANT: Bach, Alfred  
APPLICANT: Hille, Heinz  
APPLICANT: Bialojan, Siegfried  
TITLE OF INVENTION: No. 5759541el Proteins, the Preparation and Use  
TITLE OF INVENTION: Thereof  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Kell & Weinkauff  
STREET: 1101 Connecticut Avenue  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage  
COMPUTER: IBM AT-compatible, 80286 processor  
OPERATING SYSTEM: MS-DOS version 5.0  
SOFTWARE: Wordperfect version 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/684,862  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/361,705  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/966,040  
FILING DATE: 30-DEC-1992  
APPLICATION NUMBER: PCT/EP91/01361  
FILING DATE: 19-JUL-1991  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1096 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
ORIGINAL SOURCE:  
ORGANISM: Agkistrodon rhodostoma  
FEATURE:  
LOCATION: 144 to 841  
OTHER INFORMATION: the coding region shown in (2)(1x)(B)  
OTHER INFORMATION: codes for the protein of SEQ ID NO: 3  
US-08-684-862-8

Query Match 70.9%; Score 15.6; DB 1; Length 1096;  
Best Local Similarity 81.8%; Pred. No. 49;  
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 tgttgatcacagaatttgata 22  
||| | ||||| ||||| ||  
DB 112 TGTAGATCACAAGTTTCTA 91

## RESULT 15

US-08-928-383B-1/c  
; Sequence 1, Application US/08928383B  
; Patent No. 6210921  
; GENERAL INFORMATION:  
; APPLICANT: Robert W. Finberg, Jeffrey M. Bergelson,  
; APPLICANT: and Marshall S. Horwitz  
; TITLE OF INVENTION: CAR, A No. 6210921el Cocksacklevirus and Adenovirus  
; TITLE OF INVENTION: Receptor  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD, LLP  
; STREET: 28 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/928,383B  
; FILING DATE: 12-SEP-1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/026,100  
; FILING DATE: 13-SEP-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mandragouras, Amy E.  
; REGISTRATION NUMBER: 36,207  
; REFERENCE/DOCKET NUMBER: DFN-020  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617)227-7400  
; TELEFAX: (617)42-4214  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1584 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: CDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 60..1157  
; US-08-928-383B-1

Query Match 70.9%; Score 15.6; DB 4; Length 1584;  
Best Local Similarity 81.8%; Pred. No. 51;  
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
OY 1 tgttgatcacagaatttgata 22  
||| | ||||| ||||| ||  
DB 371 TGATGATCACCAGATTGAGA 350

Search completed: December 15, 2001, 03:22:46  
Job time: 7561 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 02:33:32 ; Search time 4441.54 Seconds  
(without alignments)  
53.226 Million cell updates/sec

Title: US-09-380-826a-2

Perfect score: 22  
Sequence: 1 tcttgatcacaagattgata 22

Scoring table: IDENTITY\_NUC  
Gapop 10.0 ; Gapext 1.0

Searched: 11351937 seqs, 537289281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: em\_estfun:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estom:\*  
5: em\_estcpl:\*  
6: em\_estbda:\*  
7: em\_estlro:\*  
8: em\_estlov:\*  
9: em\_hlc:\*  
10: qb\_estl:\*  
11: qb\_estl2:\*  
12: qb\_hlc:\*  
13: qb\_gss:\*  
14: em\_gss\_fun:\*  
15: em\_gss\_hum:\*  
16: em\_gss\_inv:\*  
17: em\_gss\_pln:\*  
18: em\_gss\_pro:\*  
19: em\_gss\_rtd:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	18.8	85.5	361	10	AV532040 AV532040
C 2	18.8	85.5	394	11	226559 ATG51651 Ve
C 3	18.8	85.5	490	10	AB038725 AB038725
C 4	18.8	85.5	527	10	AV520789 AV520789
C 5	18.8	85.5	530	10	AV520789 AV520789
C 6	18.4	83.6	608	11	BE977618 bs66h08.y
C 7	18.4	83.6	934	13	CNS02BWA
C 8	17.8	80.9	600	11	BF006444 EST434942
C 9	17.8	80.9	633	11	BF006254 EST434752
C 10	17.4	79.1	286	10	BR288727 BR288727
C 11	17.4	79.1	407	10	BE581720 Kq51c07.y
C 12	17.4	79.1	529	13	AQ497779 HS_5066_B

C 13	17.4	79.1	646	10	AW695022 NF082E04S
C 14	17.4	79.1	671	11	BC448853 NF003C10I
C 15	17.4	79.1	762	10	BE283043 601101323
C 16	17.4	79.1	837	10	BE642809 Crl2_7_B0
C 17	17.4	79.1	1101	13	CNS00D26
C 18	17.2	78.2	151	10	BE068225 MR4-B7036
C 19	17.2	78.2	425	10	A1049911 an30h03.x
C 20	17.2	78.2	431	13	A0595814 HS_2132_B
C 21	17.2	78.2	432	13	A0221715 HS_2010_A
C 22	17.2	78.2	437	11	W43212 22591_Lambda
C 23	17.2	78.2	439	13	AQ796989 nbx0071H
C 24	17.2	78.2	500	10	A1733664 an30h03.x
C 25	17.2	78.2	524	13	B62582 T22F18TR_TA
C 26	17.2	78.2	531	13	A0223477 HS_2003_B
C 27	17.2	78.2	536	10	AM034253 EST277824
C 28	17.2	78.2	546	10	A1779714 EST260593
C 29	17.2	78.2	579	13	B67515 T22M10TR_TA
C 30	17.2	78.2	602	10	A1777095 EST258060
C 31	17.2	78.2	614	10	BE187570 EST336131
C 32	17.2	78.2	619	10	AM963686 EST375759
C 33	17.2	78.2	638	11	BF346432 602020273
C 34	17.2	78.2	680	13	B57784 C1T-HSP-201
C 35	17.2	78.2	698	13	A2193715 SP_1023_B
C 36	17.2	78.2	773	13	CNS07DUS
C 37	17.2	78.2	888	13	CNS079F4
C 38	17.2	78.2	904	13	CNS029DN
C 39	17.2	78.2	949	13	CNS04HNP
C 40	17	77.3	373	11	C70416
C 41	17	77.3	648	10	A0200312
C 42	16.8	76.4	238	10	AV313332
C 43	16.8	76.4	260	10	BE118999 UI-R-CAO-
C 44	16.8	76.4	296	11	T20362
C 45	16.8	76.4	308	11	BI398487 952002A03

## ALIGNMENTS

RESULT 1  
LOCUS AV532040 361 bp mRNA EST 01-SEP-2000  
DEFINITION AV532040 Arabidopsis thaliana flower buds Columbia Arabidopsis  
thaliana cDNA clone FB034f09F 3', mRNA sequence.  
ACCESSION AV532040  
VERSION AV532040.1 GI:8692323  
KEYWORDS EST.  
SOURCE thale cress.  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
1 (bases 1 to 361)  
Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.  
A large scale analysis of cDNA in Arabidopsis thaliana: Generation  
of 12,028 non-redundant expressed sequence tags from normalized and  
size-selected cDNA libraries  
JOURNAL DNA Res. 7, 175-180 (2000)  
MEDLINE 20363093  
COMMENT Contact: Erika Asemitzu  
The First Laboratory for Plant Gene Research  
Kazusa DNA Research Institute  
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan  
Email: asamizuekazusa.or.jp URL: http://www.kazusa.or.jp/en/plant/.  
FEATURES  
source  
1..361  
/organism="Arabidopsis thaliana"  
/strain="Columbia"  
/db\_xref="taxon:3702"  
/clone="FB034f09F"  
/tissue\_type="flower buds"  
/note="Vector: pBluescriptII SK-; Site\_1: EcoRI; Site\_2:  
XhoI"

BASE COUNT 113 a 73 c 61 g 114 t  
 ORIGIN

Query Match 85.5%; Score 18.8; DB 10; Length 361;  
 Best Local Similarity 90.9%; Pred. No. 99;  
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 ttttgatcacagaatttgata 22  
 ||||| ||||| ||||| ||||| |||||  
 Db 276 TGTGTGTTCCACAGATTGACA 255

RESULT 2  
 LOCUS 226559/c 394 bp mRNA EST 31-MAY-1995  
 DEFINITION ATTS1651 Versaillais-VC Arabidopsis thaliana cDNA clone VCVDH08 3'  
 ACCESSION 226559  
 VERSION 226559.1 GI:404227  
 KEYWORDS EST.  
 SOURCE thale cress.  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 1 (bases 1 to 394)  
 CNRS.  
 The Arabidopsis thaliana transcribed genome: the GDR cDNA program  
 UNPUBLISHED (1996)  
 CONTACT Desprez T., Amselem J., Chlapello H., Rouze P., Caboche  
 M., Hofte H.  
 INRA Versailles  
 Laboratoire de Biologie Cellulaire  
 Route de Saint-Cyr, 78026 Versailles Cedex, France  
 Email: thierry@versailles.inra.fr.  
 Location/Qualifiers

FEATURES  
 source  
 1..394  
 /organism="Arabidopsis thaliana"  
 /strain="ecotype Columbia"  
 /db\_xref="taxon:3702"  
 /clone="VCVDH08"  
 /clone\_lib="Versailles-VC"  
 /tissue\_type="whole seedlings"  
 /dev\_stage="in vitro-grown etiolated seedlings, 5 days old"  
 /note="Vector: pbluescript"

BASE COUNT 117 a 80 c 63 g 133 t 1 others

ORIGIN

Query Match 85.5%; Score 18.8; DB 11; Length 394;  
 Best Local Similarity 90.9%; Pred. No. 1e+02;  
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 ttttgatcacagaatttgata 22  
 ||||| ||||| ||||| ||||| |||||  
 Db 321 TGTGTGTTCCACAGATTGACA 300

RESULT 3  
 LOCUS AB038725/c 490 bp mRNA EST 14-NOV-2000  
 DEFINITION AB038725 Arabidopsis thaliana Above-ground organ from two to  
 six-week old plants Columbia Arabidopsis thaliana cDNA clone  
 AP230603.f 3', mRNA sequence.  
 ACCESSION AB038725  
 VERSION AB038725.1 GI:7212552  
 KEYWORDS EST.  
 SOURCE thale cress.  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 490)  
 AUTHORS Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.  
 TITLE A large scale analysis of cDNA in Arabidopsis thaliana: Generation  
 of 12,028 non-redundant expressed sequence tags from normalized and  
 size-selected cDNA libraries  
 JOURNAL DNA Res. 7, 175-180 (2000)  
 MEDLINE 20363093  
 COMMENT Contact: Erika Asamizu  
 The First Laboratory for Plant Gene Research  
 Kazusa DNA Research Institute  
 Yana 1532-3, Kisarazu, Chiba 292-0812, Japan  
 Email: asamizuk@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.  
 Location/Qualifiers

FEATURES  
 source  
 1..490  
 /organism="Arabidopsis thaliana"  
 /strain="Columbia"  
 /db\_xref="taxon:3702"  
 /clone="AP230603.f"  
 /clone\_lib="Arabidopsis thaliana Above-ground organ from  
 two to six-week old plants Columbia"  
 /tissue\_type="aboveground organ"  
 /dev\_stage="two to six-week old plants"  
 /note="Vector: pbluescriptII SK-, Site\_1: EcoRI; Site\_2:  
 XhoI"

BASE COUNT 147 a 99 c 87 g 157 t

ORIGIN

Query Match 85.5%; Score 18.8; DB 10; Length 490;  
 Best Local Similarity 90.9%; Pred. No. 1e+02;  
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 ttttgatcacagaatttgata 22  
 ||||| ||||| ||||| ||||| |||||  
 Db 303 TGTGTGTTCCACAGATTGACA 282

RESULT 4  
 LOCUS AV520789/c 527 bp mRNA EST 01-SEP-2000  
 DEFINITION AV520789 Arabidopsis thaliana aboveground organs two to six-week  
 old Arabidopsis thaliana cDNA clone AP232606f 3', mRNA sequence.  
 ACCESSION AV520789  
 VERSION AV520789.1 GI:8680316  
 KEYWORDS EST.  
 SOURCE thale cress.  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 1 (bases 1 to 527)  
 Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.  
 A large scale analysis of cDNA in Arabidopsis thaliana: Generation  
 of 12,028 non-redundant expressed sequence tags from normalized and  
 size-selected cDNA libraries  
 JOURNAL DNA Res. 7, 175-180 (2000)  
 MEDLINE 20363093  
 COMMENT Contact: Erika Asamizu  
 The First Laboratory for Plant Gene Research  
 Kazusa DNA Research Institute  
 Yana 1532-3, Kisarazu, Chiba 292-0812, Japan  
 Email: asamizuk@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.  
 Location/Qualifiers

FEATURES  
 source  
 1..527  
 /organism="Arabidopsis thaliana"  
 /strain="Columbia"  
 /db\_xref="taxon:3702"  
 /clone="AP232606f"  
 /clone\_lib="Arabidopsis thaliana aboveground organs two to  
 six-week old"  
 /tissue\_type="aboveground organs"  
 /dev\_stage="two to six-week old"  
 /note="Vector: pbluescriptII SK-, Site\_1: EcoRI; Site\_2:"



sequence.  
 accession AL190387.1 GI:7828491  
 version GSS: genome survey sequence.  
 keywords Tetraodon nigroviridis.  
 source Tetraodon nigroviridis.  
 organism Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Acanthopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Tetraodontidae; Tetraodon.

REFERENCE  
 AUTHORS 1 (bases 1 to 934)  
 Roest-Crolius, H., Jaillon, O., Dasilva, C., Fizames, C., Fisher, C., Bouneau, L., Billault, A., Queller, F., Saurin, W., Bernot, A. and Weissenbach, J.  
 Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis

TITLE Unpublished  
 JOURNAL 2 (bases 1 to 934)  
 REFERENCE Roest-Crolius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C., Bernot, A., Fizames, C., Mincker, P., Brottier, P., Queller, F., Saurin, W. and Weissenbach, J.  
 Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence

COMMENT  
 JOURNAL Direct Submission  
 REFERENCE 3 (bases 1 to 934)  
 Genoscope.  
 TITLE Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases  
 JOURNAL This sequence is a single read and was generated as part of a large  
 COMMENT scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at  
 http://www.genoscope.cns.fr/Tetraodon.  
 Location/Qualifiers  
 1. 934  
 /organism="Tetraodon nigroviridis"  
 /db\_xref="taxon:99883"  
 /clone="254C22"  
 /clone\_1lb="G"  
 /note="Genoscope sequence ID : C0AG254B11SP1-end : PUC-ori"

BASE COUNT 282 a 175 c 207 g 259 t 11 others  
 ORIGIN

Query Match 83.6%; Score 18.4; DB 13; Length 934;  
 Best Local Similarity 95.0%; Pred. No. 1.8e+02;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 ttgatcacagaattgata 22  
 ||||| ||||| ||||| |||||  
 Db 840 TTGGAACACAGATTGATA 821

RESULT 8  
 BF006444/c 600 bp mRNA EST 06-OCT-2000  
 LOCUS EST434942 DSLC Medicago truncatula cDNA clone pDSLc-41J15, mRNA  
 DEFINITION  
 ACCESSION BF006444  
 VERSION BF006444.1 GI:10706719  
 KEYWORDS EST.  
 SOURCE barrel medic.  
 ORGANISM Medicago truncatula  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eustosids I; Fabales; Fabaceae; Papilionoidae; Trifoliaceae; Medicago.  
 1 (bases 1 to 600)  
 Fedorova, M., Pierson, B. L., Samac, D. A., Gantt, J. S., Vance, C. P., Gonzales, M. B. and Ellis, L.  
 ESTs from Medicago truncatula leaves and cotyledons  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Deborah A. Samac

Department of Plant Pathology  
 University of Minnesota  
 495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA  
 Tel: 612 625 1243  
 Fax: 651 649 5058  
 Email: debbys@pucini.crl.umn.edu  
 University of Minnesota name: M275482e TIGR sequence name: M275482e  
 More information is available at:  
 http://chrysis.tamu.edu/medicago  
 Seq primer: Skmod (CTA gaa gta gat CC).  
 Location/Qualifiers  
 1. 600  
 /organism="Medicago truncatula"  
 /cultivar="genotype A17"  
 /db\_xref="taxon:3880"  
 /clone="pDSLc-41J15"  
 /clone\_1lb="DSLc"  
 /tissue\_type="leaves and cotyledons"  
 /dev\_stage="mixture of cotyledons from five days old plants and leaves obtained from two weeks old plants"  
 /lab\_host="E. coli strain SOLR"  
 /note="Vector: pluscript SK +/-; Site 1: EcoRI; Site 2: XhoI. cDNA was prepared from polyA+ enriched RNA from the mixture of cotyledons of five days old plants and leaves of two weeks old plants. The cDNA was directionally ligated into the Uni-ZAP XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-ZAP phage using Ex-Assist helper phage and propagated in SOLR cells."

BASE COUNT 127 a 120 c 132 g 221 t  
 ORIGIN

Query Match 80.9%; Score 17.8; DB 11; Length 600;  
 Best Local Similarity 90.5%; Pred. No. 3.2e+02;  
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tttgatcacagaattgata 21  
 ||||| ||||| ||||| |||||  
 Db 90 TGTGATCACAAGATTGAT 70

RESULT 9  
 BF006254/c 633 bp mRNA EST 06-OCT-2000  
 LOCUS EST434752 DSLC Medicago truncatula cDNA clone pDSLc-40G21, mRNA  
 DEFINITION  
 ACCESSION BF006254  
 VERSION BF006254.1 GI:10706529  
 KEYWORDS EST.  
 SOURCE barrel medic.  
 ORGANISM Medicago truncatula  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eustosids I; Fabales; Fabaceae; Papilionoidae; Trifoliaceae; Medicago.  
 1 (bases 1 to 633)  
 Fedorova, M., Pierson, B. L., Samac, D. A., Gantt, J. S., Vance, C. P., Gonzales, M. B. and Ellis, L.  
 ESTs from Medicago truncatula leaves and cotyledons  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Deborah A. Samac  
 Department of Plant Pathology  
 University of Minnesota  
 495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA  
 Tel: 612 625 1243  
 Fax: 651 649 5058  
 Email: debbys@pucini.crl.umn.edu  
 University of Minnesota name: M275292e TIGR sequence name: M275292e  
 More information is available at:  
 http://chrysis.tamu.edu/medicago  
 Seq primer: Skmod (CTA gaa gta gat CC).



FEATURES  
source  
1. 633  
Location/Qualifiers  
/organism="Medicago truncatula"  
/cultivar="genotype A17"  
/db\_xref="taxon:3880"  
/clone="PDSLC-40621"  
/clone\_1ib="DSLC"  
/tissue\_type="Leaves and cotyledons"  
/dev\_stage="mixture of cotyledons from five days old plants and leaves obtained from two weeks old plants"  
/lab\_host="E. coli strain SOLR"  
/note="Vector: pBluescript SK +/-. Site\_1: EcoRI, site\_2: XhoI. cDNA was prepared from polyA+ enriched RNA from the mixture of cotyledons of five days old plants and leaves of two weeks old plants. The cDNA was directionally ligated into the Uni-ZAP XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-ZAP phage using Ex-Assist helper phage and propagated in SOLR cells."

BASE COUNT  
148 a 132 c 124 g 229 t

ORIGIN

Query Match  
Best local Similarity 90.5%; Score 17.8; DB 11; Length 633;  
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 tcttgatcacagaattgat 21  
||||| ||||||| |||

Db 188 TGTGATCATCAGATTAGAT 168

RESULT 10  
BB288727/c  
LOCUS  
DEFINITION  
BB288727 286 bp mRNA  
BB288727 RIKEN full-length enriched, 2 cells egg Mus musculus cDNA  
clone B020032C12 3', mRNA sequence.

ACCESSION  
BB288727  
VERSION  
BB288727.1 GI:8989176  
KEYWORDS  
EST.  
SOURCE  
house mouse.  
ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
1 (bases 1 to 286)  
Kono, H., Aizawa, K., Akahira, S., Akiyama, J., Arikawa, T., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Hori, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Kiyosawa, H., Kojima, Y., Kondo, S., Koyu, S., Kurihara, C., Kusakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata, Y., Shigemoto, Y., Shingawa, A., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomioka, N., Toya, T., Tsunoda, Y., Watahiki, A., Watanabe, S., Yamamura, T., Yamakawa, I., Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

TITLE  
RIKEN Mouse ESTs (Kono, H., et al.)  
JOURNAL  
COMMENT  
Unpublished (2000)  
Contact: Yoshinori Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suenho-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-resgsc.riken.go.jp,  
URL: http://genome-gsc.riken.go.jp/  
Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagasaki, S., Sasaki, N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Thermolabile and thermocyclization of thermolabile enzymes by trehalose and its application for the synthesis of full length

CDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)  
Itoh, M., Kitanai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.  
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)  
Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)  
Please visit our web site (<http://genome.riken.go.jp>) for further details.

FEATURES  
source  
1. 286  
Location/Qualifiers  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="B020032C12"  
/clone\_1ib="RIKEN full-length enriched, 2 cells egg"  
/tissue\_type="egg"  
/dev\_stage="2 cells"  
/lab\_host="DH10B"  
/note="Site\_1: SalI; Site\_2: BamHI. cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGGACAGATCCAGACCTCTTTTCTTTTCTTTTNN 3']. cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence 15' GAGAGGACATCTCGAGTTATTAATTAATCCGCCGCCGCC 3'. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FIC I."

BASE COUNT  
117 a 66 c 20 g 83 t

ORIGIN

Query Match  
Best local Similarity 94.7%; Score 17.4; DB 10; Length 286;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 gtgcgacacagaattga 20  
||||| ||||||| |||

Db 169 GTTGATCACAATTTGA 151

RESULT 11  
BE581720/c  
LOCUS  
DEFINITION  
BE581720 407 bp mRNA  
BE581720 RIKEN full-length enriched, 2 cells egg Mus musculus cDNA  
clone B020032C12 3', mRNA sequence.

ACCESSION  
BE581720  
VERSION  
BE581720.1 GI:9832662  
KEYWORDS  
EST.  
SOURCE  
Strongyloides stercoralis.  
ORGANISM  
Strongyloides stercoralis.  
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Panagrolaimidae; Strongyloididae; Strongyloides.

REFERENCE  
1 (bases 1 to 407)  
McCartner, J., Clifton, S., Chiapelli, B., Pape, D., Martin, J., Wylie, T., Dante, M., Marra, M., Hillier, L., Kucaba, T., Theising, B., Bowers, Y., Gibbons, M., Ritzer, E., Bennett, J., Franklin, C., Tsagarelis, V., Runko, I., Kennedy, S., Maguire, L., Beck, C., Underwood, K., Steptoe, M., Allen, M., Person, B., Swaller, T., Harvey, N., Schurk, R., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterson, R. and Wilson, R.

TITLE  
The Washington Univ. Nematode EST Project, 1999  
JOURNAL  
COMMENT  
Unpublished (1999)  
Contact: McCartner, JP  
The Washington Univ. Nematode EST Project, 1999  
Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810

Email: estevenson.wustl.edu  
The library was constructed by Dr. Thomas Nulman and colleagues of  
NIH (nutman@nih.gov). DNA Sequencing by: Washington  
University Genome Sequencing Center St. Louis.  
High quality sequence stop: 384.  
Location/Qualifiers

## FEATURES

## SOURCE

1. 407  
/organism="Strongyloides stercoralis"  
/strain="Rhabditiform larvae obtained from gerbils"  
/db\_xref="taxon:6248"  
/clone\_lib="TBN95TM-SSR"  
/lab\_host="XL-1 Blue MRF" (Stratagene); Site: 1:  
/note="Vector: Lambda Uni-ZAP XR (Stratagene); Site: 1:  
EcoRI; Site: 2: XhoI; mRNA had been isolated from gerbils  
rhabditiform larvae which had been isolated from gerbils  
experimentally infected with larvae originally isolated  
from experimentally infected dogs. cDNA was constructed  
and, using adaptors, was cloned unidirectionally into the  
vector from the EcoRI site to the XhoI site. The library  
has an unamplified titer of 1 x 10<sup>5</sup> pfu/ml and an  
amplified, undiluted titer of 9 x 10<sup>5</sup> pfu/ml. The  
average insert size of the unamplified library is 675 bp  
(range, 100-1700)."

BASE COUNT 152 a 63 c 74 g 118 t  
ORIGIN

Query Match 79.1%; Score 17.4; DB 10: Length 407;  
Best Local Similarity 94.7%; Pred. No. 4.8e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 ttggtcacaaagatttgata 22  
1 |||||  
Db 334 TTGATCACAAGATTGATA 316

## RESULT 12

LOCUS A0497779 529 bp DNA GSS 28-APR-1999  
DEFINITION H5.5066\_B1.E05.T7A RPCI-11 Human Male BAC Library Homo sapiens  
genomic clone Plate=642 Col=9 Row=J, DNA sequence.

ACCESSION A0497779  
KEYWORDS A0497779.1 GI:4697902  
GSS.

SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 529)  
Mahalir, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,  
Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and  
Hood, L.  
Sequence-tagged connectors: A sequence approach to mapping and  
scanning the human genome  
Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)

TITLE High Throughput Sequencing Center  
JOURNAL Contact: Mahalir GG, Wallace JC, Hood L  
MEDLINE University of Washington  
COMMENT 401 Queen Anne Avenue North, Seattle, WA 98109, USA  
Tel: (206) 616-3618  
Fax: (206) 616-3887  
Email: jwallace@u.washington.edu  
Library are derived from the human BAC library RPCI-11. For BAC  
clones availability, please contact Pieter de Jong  
(pieter@dejong.med.bufileo.edu). Clones may be purchased from  
BACPAC Resources (http://bacpac.med.bufileo.edu/ordering.bac.htm)  
or from Research Genetics (inf@resgen.com). BAC end Web Server:  
http://www.htsc.washington.edu  
Plate: 642 row: J column: 9

Seq primer: T7  
Class: BAC ends  
High quality sequence stop: 529.  
Location/Qualifiers

## FEATURES

## SOURCE

1. 529  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="Plate=642 Col=9 Row=J"  
/clone\_lib="RPCI-11 Human Male BAC Library"  
/sex="Male"  
/note="Vector: pBAC3.6; Site: 1: EcoRI; Site: 2: EcoRI;  
Male blood DNA was isolated from one randomly chosen donor  
and partially digested with a combination of EcoRI and  
EcoRI Methylase. Size selected DNA was cloned into the  
pBAC3.6 vector at EcoRI sites"  
pBAC3.6 vector at EcoRI sites" 11 others

BASE COUNT 145 a 99 c 109 g 165 t  
ORIGIN

Query Match 79.1%; Score 17.4; DB 13: Length 529;  
Best Local Similarity 94.7%; Pred. No. 4.8e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 ttggtcacaaagatttgat 21  
1 |||||  
Db 403 TTGATCACAAGATTGAT 421

## RESULT 13

LOCUS AM695022/c 646 bp mRNA EST 21-DEC-2000  
DEFINITION NF082E04ST1P1034 Developing stem Medicago truncatula cDNA clone  
NF082E04ST 5', mRNA sequence.

ACCESSION AM695022  
KEYWORDS AM695022.2 GI:11957200  
EST.

SOURCE barrel medic.  
ORGANISM Medicago truncatula  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;  
Medicago.  
1 (bases 1 to 646)  
He, X.-Z., Shadle, G., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell  
C.J., Flores, H.R., Inman, J.T., Weller, J.W., May, G.D. and Dixon  
R.A.

REFERENCE 1 (bases 1 to 646)  
He, X.-Z., Shadle, G., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell  
C.J., Flores, H.R., Inman, J.T., Weller, J.W., May, G.D. and Dixon  
R.A.  
Expressed Sequence Tags from the Samuel Roberts Noble Foundation  
Medicago truncatula stem library  
Unpublished (2000)  
On Apr 14, 2000 this sequence version replaced gi:7569784.

TITLE Plant Biology Division  
JOURNAL The Samuel Roberts Noble Foundation  
COMMENT 2510 Sam Noble Parkway, Ardmore, OK 73402, USA  
Tel: 580 221 7302  
Fax: 580 221 7380  
Email: radixon@noble.org  
Insert Length: 674 Std Error: 0.00  
Plate: 082 row: E column: 04  
Seq primer: TCACACAGGAACACGTATGAC.  
Location/Qualifiers

## FEATURES

## SOURCE

1. 646  
/organism="Medicago truncatula"  
/db\_xref="taxon:3880"  
/clone\_lib="NF082E04ST"  
/clone\_lib="Developing stem"  
/issue\_type="stem"  
/dev\_stage="Pooled developmental"  
/note="Vector: Lambda Zap; Contains a mixture of  
intermodal stem segments"

## BASE COUNT

180 a 142 c 129 g 195 t  
ORIGIN



---

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 05:42:56 : Search time 7962.77 Seconds  
(Without alignments)  
15.407 Million cell updates/sec

Title: US-09-380-826a-5  
Perfect score: 7  
Sequence: 1 ttgata 7

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 17159718 seqs, 8763200856 residues

Total number of hits satisfying chosen parameters: 34319436

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_NA\_Main:\*

1:	/cgn2_6/prodata/2/pna/PCYUS_COMB.seq:*
2:	/cgn2_6/prodata/2/pna/US06_COMB.seq:*
3:	/cgn2_6/prodata/2/pna/US07_COMB.seq:*
4:	/cgn2_6/prodata/2/pna/US080_COMB.seq:*
5:	/cgn2_6/prodata/2/pna/US081_COMB.seq:*
6:	/cgn2_6/prodata/2/pna/US082_COMB.seq:*
7:	/cgn2_6/prodata/2/pna/US083_COMB.seq:*
8:	/cgn2_6/prodata/2/pna/US084_COMB.seq:*
9:	/cgn2_6/prodata/2/pna/US085_COMB.seq:*
10:	/cgn2_6/prodata/2/pna/US086_COMB.seq:*
11:	/cgn2_6/prodata/2/pna/US087_COMB.seq:*
12:	/cgn2_6/prodata/2/pna/US088_COMB.seq:*
13:	/cgn2_6/prodata/2/pna/US089_COMB.seq:*
14:	/cgn2_6/prodata/2/pna/US090_COMB.seq:*
15:	/cgn2_6/prodata/2/pna/US091_COMB.seq:*
16:	/cgn2_6/prodata/2/pna/US092_COMB.seq:*
17:	/cgn2_6/prodata/2/pna/US093_COMB.seq:*
18:	/cgn2_6/prodata/2/pna/US094_COMB.seq:*
19:	/cgn2_6/prodata/2/pna/US095A_COMB.seq:*
20:	/cgn2_6/prodata/2/pna/US095B_COMB.seq:*
21:	/cgn2_6/prodata/2/pna/US095C_COMB.seq:*
22:	/cgn2_6/prodata/2/pna/US095D_COMB.seq:*
23:	/cgn2_6/prodata/2/pna/US096A_COMB.seq:*
24:	/cgn2_6/prodata/2/pna/US096B_COMB.seq:*
25:	/cgn2_6/prodata/2/pna/US096C_COMB.seq:*
26:	/cgn2_6/prodata/2/pna/US096D_COMB.seq:*
27:	/cgn2_6/prodata/2/pna/US096E_COMB.seq:*
28:	/cgn2_6/prodata/2/pna/US096F_COMB.seq:*
29:	/cgn2_6/prodata/2/pna/US097A_COMB.seq:*
30:	/cgn2_6/prodata/2/pna/US097C_COMB.seq:*
31:	/cgn2_6/prodata/2/pna/US099_COMB.seq:*
32:	/cgn2_6/prodata/2/pna/US099A_COMB.seq:*
33:	/cgn2_6/prodata/2/pna/US6000_COMB.seq:*
34:	/cgn2_6/prodata/2/pna/US6001_COMB.seq:*
35:	/cgn2_6/prodata/2/pna/US6002_COMB.seq:*
36:	/cgn2_6/prodata/2/pna/US6003_COMB.seq:*
37:	/cgn2_6/prodata/2/pna/US6004_COMB.seq:*
38:	/cgn2_6/prodata/2/pna/US6005_COMB.seq:*
39:	/cgn2_6/prodata/2/pna/US6006_COMB.seq:*
40:	/cgn2_6/prodata/2/pna/US6007_COMB.seq:*
41:	/cgn2_6/prodata/2/pna/US6008_COMB.seq:*
42:	/cgn2_6/prodata/2/pna/US6009_COMB.seq:*
43:	/cgn2_6/prodata/2/pna/US6010_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7	100.0	7	17	US-09-380-826a-5
2	7	100.0	10	17	PCT-US99-13800-1005
3	7	100.0	10	17	US-09-335-032-1287
4	7	100.0	10	17	US-09-335-032-1793
5	7	100.0	10	17	US-09-335-032-4831
6	7	100.0	10	17	US-09-335-032-6568
7	7	100.0	10	17	US-09-335-032-7108
8	7	100.0	10	17	US-09-335-032-7128
9	7	100.0	10	17	US-09-335-032-7153
10	7	100.0	10	17	US-09-335-032-7517
11	7	100.0	10	17	US-09-335-032-7816
12	7	100.0	10	17	US-09-335-032-8062
13	7	100.0	10	17	US-09-335-032-10702
14	7	100.0	10	17	US-09-335-032-11663
15	7	100.0	10	17	US-09-335-032-11665
16	7	100.0	10	17	US-09-335-032-11665
17	7	100.0	10	17	US-09-335-032-12098
18	7	100.0	10	17	US-09-335-032-12170
19	7	100.0	10	17	US-09-335-032-12170
20	7	100.0	10	17	US-09-335-032-12170
21	7	100.0	10	17	US-09-335-032-12170
22	7	100.0	10	17	US-09-335-032-12170
23	7	100.0	10	17	US-09-335-032-12170
24	7	100.0	10	17	US-09-335-032-12170
25	7	100.0	10	17	US-09-335-032-12170
26	7	100.0	10	17	US-09-335-032-12170
27	7	100.0	10	17	US-09-335-032-12170
28	7	100.0	10	17	US-09-335-032-12170
29	7	100.0	10	17	US-09-335-032-12170
30	7	100.0	10	17	US-09-335-032-12170
31	7	100.0	10	17	US-09-335-032-12170
32	7	100.0	10	17	US-09-335-032-12170
33	7	100.0	10	17	US-09-335-032-12170
34	7	100.0	10	17	US-09-335-032-12170
35	7	100.0	10	17	US-09-335-032-12170
36	7	100.0	10	17	US-09-335-032-12170
37	7	100.0	10	17	US-09-335-032-12170
38	7	100.0	10	17	US-09-335-032-12170
39	7	100.0	10	17	US-09-335-032-12170
40	7	100.0	10	17	US-09-335-032-12170
41	7	100.0	10	17	US-09-335-032-12170
42	7	100.0	10	17	US-09-335-032-12170
43	7	100.0	10	17	US-09-335-032-12170

c 42 7 100.0 15 9 US-08-584-040-8444  
c 43 7 100.0 15 9 US-08-585-684-113  
c 44 7 100.0 15 9 US-08-585-684-114  
c 45 7 100.0 15 28 US-09-708-690-4100

Sequence 8444, Ap  
Sequence 113, App  
Sequence 114, App  
Sequence 4100, Ap

## ALIGNMENTS

RESULT 1  
US-09-380-826a-5  
Sequence 5, Application US/09380826A  
GENERAL INFORMATION:  
APPLICANT: Chappel, Rod  
TITLE OF INVENTION: LEPTOSPIRA PATHOGENS  
FILE REFERENCE: DAVI79.001APC  
CURRENT APPLICATION NUMBER: US/09/380,826A  
CURRENT FILING DATE: 1999-11-22  
PRIOR APPLICATION NUMBER: PCT/AU98/00145  
PRIOR FILING DATE: 1998-03-06  
PRIOR APPLICATION NUMBER: AU P05494/97  
PRIOR FILING DATE: 1997-08-07  
NUMBER OF SEQ ID NOS: 26  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 5  
LENGTH: 7  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Primer for L. fainei.  
US-09-380-826a-5

Query Match 100.0%; Score 7; DB 17; Length 7;  
Best Local Similarity 100.0%; Pred. No. 2.5e+09;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tttagata 7  
Db 1 tttagata 7

RESULT 2  
PCT-US99-13800-1005/c  
Sequence 1005, Application PC/RTUS9913800  
GENERAL INFORMATION:  
APPLICANT: Genzyme Corporation  
APPLICANT: Roberts, Bruce L.  
APPLICANT: Shankara, Srinivas  
TITLE OF INVENTION: PREPARATION AND USE OF SUPERIOR VACCINES  
FILE REFERENCE: 68126881206940  
CURRENT APPLICATION NUMBER: PCT/US99/13800  
CURRENT FILING DATE: 1999-06-18  
EARLIER APPLICATION NUMBER: 60/090,039  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/090,040  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/090,041  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/089,853  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/089,997  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/090,079  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/090,035  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/089,993  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/089,992  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/090,072  
EARLIER FILING DATE: 1998-06-19

EARLIER APPLICATION NUMBER: 60/089,878  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/089,991  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/090,000  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/090,048  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/089,999  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/090,043  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/090,042  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/090,036  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/090,044  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/089,844  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/090,080  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/089,833  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/089,994  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/090,077  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/090,078  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/090,047  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/090,076  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/090,045  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/111,715  
EARLIER FILING DATE: 1998-12-08  
NUMBER OF SEQ ID NOS: 2138  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 1005  
LENGTH: 10  
TYPE: DNA  
ORGANISM: Homo sapiens  
PCT-US99-13800-1005  
Query Match 100.0%; Score 7; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 3.3e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 tttagata 7  
Db 9 TTTCGATA 3  
RESULT 3  
US-09-335-032-1287/c  
Sequence 1287, Application US/09335032  
GENERAL INFORMATION:  
APPLICANT: Velculescu, Victor  
APPLICANT: Vogelstein, Bert  
APPLICANT: Kinzler, Kenneth  
TITLE OF INVENTION: Characterization of the Yeast  
FILE REFERENCE: 01107.78572  
CURRENT APPLICATION NUMBER: US/09/335,032  
CURRENT FILING DATE: 1999-06-16  
PRIOR APPLICATION NUMBER: US 60/035,917  
PRIOR FILING DATE: 1997-01-23  
PRIOR APPLICATION NUMBER: US 09/012,031  
PRIOR FILING DATE: 1998-01-22  
NUMBER OF SEQ ID NOS: 12219

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1287
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-1287
```

```
Query Match          100.0%; Score 7; DB 17; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 ttgata 7
        |||
        8 TTTGATA 2
```

```
RESULT 4
US-09-335-032-1793/c
; Sequence 1793, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335.032
; CURRENT FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1793
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-1793
```

```
Query Match          100.0%; Score 7; DB 17; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 ttgata 7
        |||
        8 TTTGATA 2
```

```
RESULT 5
US-09-335-032-4831/c
; Sequence 4831, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335.032
; CURRENT FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4831
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
```

```
US-09-335-032-4831
```

```
Query Match          100.0%; Score 7; DB 17; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 ttgata 7
        |||
        7 TTTGATA 1
```

```
RESULT 6
US-09-335-032-6568/c
; Sequence 6568, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335.032
; CURRENT FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6568
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-6568
```

```
Query Match          100.0%; Score 7; DB 17; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 ttgata 7
        |||
        10 TTTGATA 4
```

```
RESULT 7
US-09-335-032-7108/c
; Sequence 7108, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335.032
; CURRENT FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7108
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-7108
```

```
Query Match          100.0%; Score 7; DB 17; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.3e+05;
```

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgtata 7  
| | | | |  
Db 10 TTTGATA 4

RESULT 8  
US-09-335-032-7128  
; Sequence 7128, Application US/09335032  
; GENERAL INFORMATION:  
; APPLICANT: Velculescu, Victor  
; APPLICANT: Vogelstein, Bert  
; APPLICANT: Kinzler, Kenneth  
; TITLE OF INVENTION: Characterization of the Yeast  
; FILE REFERENCE: 01107.78572  
; CURRENT APPLICATION NUMBER: US/09/335,032  
; CURRENT FILING DATE: 1999-06-16  
; PRIOR APPLICATION NUMBER: US 60/035,917  
; PRIOR FILING DATE: 1997-01-23  
; PRIOR APPLICATION NUMBER: US 09/012,031  
; PRIOR FILING DATE: 1998-01-22  
; NUMBER OF SEQ ID NOS: 12219  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7128  
; LENGTH: 10  
; TYPE: DNA  
; ORGANISM: Saccharomyces cerevisiae  
US-09-335-032-7128

Query Match 100.0%; Score 7; DB 17; Length 10;  
Best Local Similarity 100.0%; Pred. No. 3.3e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgtata 7  
| | | | |  
Db 2 ttgtata 8

RESULT 9  
US-09-335-032-7153  
; Sequence 7153, Application US/09335032  
; GENERAL INFORMATION:  
; APPLICANT: Velculescu, Victor  
; APPLICANT: Vogelstein, Bert  
; APPLICANT: Kinzler, Kenneth  
; TITLE OF INVENTION: Characterization of the Yeast  
; FILE REFERENCE: 01107.78572  
; CURRENT APPLICATION NUMBER: US/09/335,032  
; CURRENT FILING DATE: 1999-06-16  
; PRIOR APPLICATION NUMBER: US 60/035,917  
; PRIOR FILING DATE: 1997-01-23  
; PRIOR APPLICATION NUMBER: US 09/012,031  
; PRIOR FILING DATE: 1998-01-22  
; NUMBER OF SEQ ID NOS: 12219  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7153  
; LENGTH: 10  
; TYPE: DNA  
; ORGANISM: Saccharomyces cerevisiae  
US-09-335-032-7153

Query Match 100.0%; Score 7; DB 17; Length 10;  
Best Local Similarity 100.0%; Pred. No. 3.3e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgtata 7  
| | | | |  
Db 3 ttgtata 9

RESULT 10  
US-09-335-032-7517  
; Sequence 7517, Application US/09335032  
; GENERAL INFORMATION:  
; APPLICANT: Velculescu, Victor  
; APPLICANT: Vogelstein, Bert  
; APPLICANT: Kinzler, Kenneth  
; TITLE OF INVENTION: Characterization of the Yeast  
; FILE REFERENCE: 01107.78572  
; CURRENT APPLICATION NUMBER: US/09/335,032  
; CURRENT FILING DATE: 1999-06-16  
; PRIOR APPLICATION NUMBER: US 60/035,917  
; PRIOR FILING DATE: 1997-01-23  
; PRIOR APPLICATION NUMBER: US 09/012,031  
; PRIOR FILING DATE: 1998-01-22  
; NUMBER OF SEQ ID NOS: 12219  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7517  
; LENGTH: 10  
; TYPE: DNA  
; ORGANISM: Saccharomyces cerevisiae  
US-09-335-032-7517

Query Match 100.0%; Score 7; DB 17; Length 10;  
Best Local Similarity 100.0%; Pred. No. 3.3e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgtata 7  
| | | | |  
Db 3 ttgtata 9

RESULT 11  
US-09-335-032-7816  
; Sequence 7816, Application US/09335032  
; GENERAL INFORMATION:  
; APPLICANT: Velculescu, Victor  
; APPLICANT: Vogelstein, Bert  
; APPLICANT: Kinzler, Kenneth  
; TITLE OF INVENTION: Characterization of the Yeast  
; FILE REFERENCE: 01107.78572  
; CURRENT APPLICATION NUMBER: US/09/335,032  
; CURRENT FILING DATE: 1999-06-16  
; PRIOR APPLICATION NUMBER: US 60/035,917  
; PRIOR FILING DATE: 1997-01-23  
; PRIOR APPLICATION NUMBER: US 09/012,031  
; PRIOR FILING DATE: 1998-01-22  
; NUMBER OF SEQ ID NOS: 12219  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7816  
; LENGTH: 10  
; TYPE: DNA  
; ORGANISM: Saccharomyces cerevisiae  
US-09-335-032-7816

Query Match 100.0%; Score 7; DB 17; Length 10;  
Best Local Similarity 100.0%; Pred. No. 3.3e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttgtata 7  
| | | | |  
Db 3 ttgtata 9

RESULT 12  
US-09-335-032-8062/c  
; Sequence 8062, Application US/09335032



```
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335,032
; CURRENT FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8062
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-8062
```

```
Query Match          100.0%; Score 7; DB 17; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 ttgata 7
    |||||
Db 8 TTTGATA 2
```

```
RESULT 13
US-09-335-032-10702/c
; Sequence 10702, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335,032
; CURRENT FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10702
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-10702
```

```
Query Match          100.0%; Score 7; DB 17; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 ttgata 7
    |||||
Db 7 TTTGATA 1
```

```
RESULT 14
US-09-335-032-11663/c
; Sequence 11663, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
```

```
; TITLE OF INVENTION: Transcriptome
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335,032
; CURRENT FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11663
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-11663
```

```
Query Match          100.0%; Score 7; DB 17; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 ttgata 7
    |||||
Db 8 TTTGATA 2
```

```
RESULT 15
US-09-335-032-11664/c
; Sequence 11664, Application US/09335032
; GENERAL INFORMATION:
; APPLICANT: Velculescu, Victor
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: Characterization of the Yeast
; FILE REFERENCE: 01107.78572
; CURRENT APPLICATION NUMBER: US/09/335,032
; CURRENT FILING DATE: 1999-06-16
; PRIOR APPLICATION NUMBER: US 60/035,917
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 09/012,031
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 12219
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11664
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-335-032-11664
```

```
Query Match          100.0%; Score 7; DB 17; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 ttgata 7
    |||||
Db 8 TTTGATA 2
```

```
Search completed: December 15, 2001, 05:42:57
Job time: 15822 sec
```



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 05:58:20 : Search time 474.83 seconds  
(without alignments)  
21.200 Million cell updates/sec

Title: US-09-380-826a-5  
Perfect score: 7  
Sequence: 1 ttgata 7

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 973175 seqs, 719035916 residues  
Total number of hits satisfying chosen parameters: 1946350

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_NA\_New:\*  
1: /cgn2\_6/ptodata/1/pna/PCR\_NEW\_COMB.seq:\*  
2: /cgn2\_6/ptodata/1/pna/US06\_NEW\_COMB.seq:\*  
3: /cgn2\_6/ptodata/1/pna/US07\_NEW\_COMB.seq:\*  
4: /cgn2\_6/ptodata/1/pna/US08\_NEW\_COMB.seq:\*  
5: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:\*  
6: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq1:\*  
7: /cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:\*  
8: /cgn2\_6/ptodata/1/pna/US60\_NEW\_COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	7	100.0	15	6	US-09-685-664B-4100 Sequence 4100, Ap
C 2	7	100.0	15	6	US-09-870-161-4100 Sequence 4100, Ap
C 3	7	100.0	16	6	US-09-670-607-1266 Sequence 1266, Ap
C 4	7	100.0	16	6	US-09-870-161-5694 Sequence 5694, Ap
C 5	7	100.0	16	6	US-09-870-161-5978 Sequence 5978, Ap
C 6	7	100.0	17	6	US-09-685-664B-116 Sequence 116, App
C 7	7	100.0	17	6	US-09-685-664B-117 Sequence 117, App
C 8	7	100.0	17	6	US-09-685-664B-118 Sequence 118, App
C 9	7	100.0	17	6	US-09-685-664B-882 Sequence 882, App
C 10	7	100.0	17	6	US-09-685-664B-883 Sequence 883, App
C 11	7	100.0	17	6	US-09-685-664B-2417 Sequence 2417, Ap
C 12	7	100.0	17	6	US-09-685-664B-2418 Sequence 2418, Ap
C 13	7	100.0	17	6	US-09-685-664B-3108 Sequence 3108, Ap
C 14	7	100.0	17	6	US-09-685-664B-3109 Sequence 3109, Ap
C 15	7	100.0	17	6	US-09-685-664B-3110 Sequence 3110, Ap
C 16	7	100.0	17	6	US-09-685-664B-3196 Sequence 3196, Ap
C 17	7	100.0	17	6	US-09-685-664B-3197 Sequence 3197, Ap
C 18	7	100.0	17	6	US-09-685-664B-3198 Sequence 3198, Ap
C 19	7	100.0	17	6	US-09-670-607-34 Sequence 34, App1
C 20	7	100.0	17	6	US-09-670-607-35 Sequence 35, App1
C 21	7	100.0	17	6	US-09-670-607-36 Sequence 36, App1
C 22	7	100.0	17	6	US-09-670-607-37 Sequence 37, App1
C 23	7	100.0	17	6	US-09-670-607-273 Sequence 273, App
C 24	7	100.0	17	6	US-09-670-607-274 Sequence 274, App
C 25	7	100.0	17	6	US-09-670-607-275 Sequence 275, App

C 26	7	100.0	17	6	US-09-670-607-276 Sequence 276, App
C 27	7	100.0	17	6	US-09-670-607-971 Sequence 971, App
C 28	7	100.0	17	6	US-09-670-607-972 Sequence 972, App
C 29	7	100.0	17	6	US-09-670-607-1467 Sequence 1467, App
C 30	7	100.0	17	6	US-09-670-607-1468 Sequence 1468, App
C 31	7	100.0	17	6	US-09-870-161-1116 Sequence 1116, App
C 32	7	100.0	17	6	US-09-870-161-1117 Sequence 117, App
C 33	7	100.0	17	6	US-09-870-161-1118 Sequence 118, App
C 34	7	100.0	17	6	US-09-870-161-882 Sequence 882, App
C 35	7	100.0	17	6	US-09-870-161-883 Sequence 883, App
C 36	7	100.0	17	6	US-09-870-161-2417 Sequence 2417, App
C 37	7	100.0	17	6	US-09-870-161-2418 Sequence 2418, App
C 38	7	100.0	17	6	US-09-870-161-3108 Sequence 3108, App
C 39	7	100.0	17	6	US-09-870-161-3109 Sequence 3109, App
C 40	7	100.0	17	6	US-09-870-161-3110 Sequence 3110, App
C 41	7	100.0	17	6	US-09-870-161-3196 Sequence 3196, App
C 42	7	100.0	17	6	US-09-870-161-3197 Sequence 3197, App
C 43	7	100.0	17	6	US-09-870-161-3198 Sequence 3198, App
C 44	7	100.0	17	6	US-09-870-161-4320 Sequence 4320, App
C 45	7	100.0	17	6	US-09-870-161-4321 Sequence 4321, App

## ALIGNMENTS

```
RESULT 1
US-09-685-664B-4100/c
; Sequence 4100, Application US/09685664B
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: Pavco, Pam
; APPLICANT: McSwiggen, Jim
; APPLICANT: Stinchcomb, Dan
; APPLICANT: Escobedo, Jaime
; TITLE OF INVENTION: Method and Reagent for Treatment of Diseases or Conditions Rel
; FILE REFERENCE: MBH00-876-K (400/021)
; CURRENT APPLICATION NUMBER: US/09/685,664B
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: US 60/005,974
; PRIOR FILING DATE: 1995-10-26
; PRIOR APPLICATION NUMBER: US 08/584,040
; PRIOR FILING DATE: 1996-01-08
; PRIOR APPLICATION NUMBER: US 09/371,772
; PRIOR FILING DATE: 1999-08-10
; NUMBER OF SEQ ID NOS: 8231
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4100
; LENGTH: 15
; TYPE: RNA
; ORGANISM: Homo sapiens
US-09-685-664B-4100

Query Match 100.0%; Score 7; DB 6; Length 15;
Best Local Similarity 100.0%; Pred. No. 7.1e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ttgata 7
Db 13 TTTGATA 7

RESULT 2
US-09-870-161-4100/c
; Sequence 4100, Application US/09870161
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: Pavco, Pam
; APPLICANT: McSwiggen, Jim
; APPLICANT: Stinchcomb, Dan
; APPLICANT: Escobedo, Jaime
; TITLE OF INVENTION: Method and Reagent for the Treatment of Diseases or Conditions
```

```
;; TITLE OF INVENTION: Levels of Vascular Endothelial Growth Factor Receptor
;; FILE REFERENCE: MBHB00-876-M (400/026)
;; CURRENT APPLICATION NUMBER: US/09/870,161
;; CURRENT FILING DATE: 2001-08-27
;; NUMBER OF SEQ ID NOS: 20821
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 4100
;; LENGTH: 15
;; TYPE: RNA
;; ORGANISM: Homo sapiens
US-09-870-161-4100
```

```
Query Match          100.0%; Score 7; DB 6; Length 15;
Best Local Similarity 100.0%; Pred. No. 7.1e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 ttgata 7
        |||||
DB      13 TTGATA 7
```

```
RESULT      3
US-09-670-607-1266
; Sequence 1266, Application US/09670607
; GENERAL INFORMATION:
; APPLICANT: Ribozyne Pharmaceuticals, Inc.
; APPLICANT: Usman, Nassim
; APPLICANT: MCSwigen, Jim
; TITLE OF INVENTION: Method and Reagent for the Inhibition of Protein-Tyrosine Phosphatase
; FILE REFERENCE: MBHB00-836-A (237/194)
; CURRENT APPLICATION NUMBER: US/09/670,607
; CURRENT FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 4262
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1266
; LENGTH: 16
; TYPE: RNA
; ORGANISM: Homo sapiens
US-09-670-607-1266
```

```
Query Match          100.0%; Score 7; DB 6; Length 16;
Best Local Similarity 42.9%; Pred. No. 7.1e+04;
Matches 3; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 ttgata 7
        ::|||
DB      8 uuugaaa 14
```

```
RESULT      4
US-09-870-161-5694/C
; Sequence 5694, Application US/09870161
; GENERAL INFORMATION:
; APPLICANT: Ribozyne Pharmaceuticals, Inc.
; APPLICANT: Pavco, Pam
; APPLICANT: MCSwigen, Jim
; APPLICANT: Stinchcomb, Dan
; APPLICANT: Escobedo, Jaime
; TITLE OF INVENTION: Method and Reagent for the Treatment of Diseases or Conditions Related to Vascular Endothelial Growth Factor Receptor
; FILE REFERENCE: MBHB00-876-M (400/026)
; CURRENT APPLICATION NUMBER: US/09/870,161
; CURRENT FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 20821
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5694
; LENGTH: 16
; TYPE: RNA
; ORGANISM: Homo sapiens
US-09-870-161-5694
```

```
Query Match          100.0%; Score 7; DB 6; Length 16;
Best Local Similarity 100.0%; Pred. No. 7.1e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 ttgata 7
        |||||
DB      9 TTGATA 3
```

```
RESULT      5
US-09-870-161-5978/C
; Sequence 5978, Application US/09870161
; GENERAL INFORMATION:
; APPLICANT: Ribozyne Pharmaceuticals, Inc.
; APPLICANT: Pavco, Pam
; APPLICANT: MCSwigen, Jim
; APPLICANT: Stinchcomb, Dan
; APPLICANT: Escobedo, Jaime
; TITLE OF INVENTION: Method and Reagent for the Treatment of Diseases or Conditions Related to Vascular Endothelial Growth Factor Receptor
; FILE REFERENCE: MBHB00-876-M (400/026)
; CURRENT APPLICATION NUMBER: US/09/870,161
; CURRENT FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 20821
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5978
; LENGTH: 16
; TYPE: RNA
; ORGANISM: Homo sapiens
US-09-870-161-5978
```

```
Query Match          100.0%; Score 7; DB 6; Length 16;
Best Local Similarity 100.0%; Pred. No. 7.1e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 ttgata 7
        |||||
DB      8 TTGATA 2
```

```
RESULT      6
US-09-685-664B-116/C
; Sequence 116, Application US/09685664B
; GENERAL INFORMATION:
; APPLICANT: Ribozyne Pharmaceuticals, Inc.
; APPLICANT: Pavco, Pam
; APPLICANT: MCSwigen, Jim
; APPLICANT: Stinchcomb, Dan
; APPLICANT: Escobedo, Jaime
; TITLE OF INVENTION: Method and Reagent for Treatment of Diseases or Conditions Related to Vascular Endothelial Growth Factor Receptor
; FILE REFERENCE: MBHB00-876-K (400/021)
; CURRENT APPLICATION NUMBER: US/09/685,664B
; CURRENT FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: US 60/005,974
; PRIOR FILING DATE: 1995-10-26
; PRIOR APPLICATION NUMBER: US 08/584,040
; PRIOR FILING DATE: 1996-01-08
; PRIOR APPLICATION NUMBER: US 09/371,772
; PRIOR FILING DATE: 1999-08-10
; NUMBER OF SEQ ID NOS: 8231
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 116
; LENGTH: 17
; TYPE: RNA
; ORGANISM: Homo sapiens
US-09-685-664B-116
```

```
Query Match          100.0%; Score 7; DB 6; Length 17;
```

Best Local Similarity 100.0%; Pred. No. 7.1e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0;  
Gaps 0;  
QY 1 ttgtgata 7  
|||||||  
DB 17 TTTGATA 11

RESULT 7  
US-09-685-664B-117/C

; Sequence 117, Application US/09685664B  
; GENERAL INFORMATION:  
; APPLICANT: Ribozyme Pharmaceuticals, Inc.  
; APPLICANT: Pavco, Pam  
; APPLICANT: MCSwigen, Jim  
; APPLICANT: Stinchcomb, Dan  
; APPLICANT: Escobedo, Jaime  
; TITLE OF INVENTION: Method and Reagent for Treatment of Diseases or Conditions Related to  
; TITLE OF INVENTION: Levels of Vascular Endothelial Growth Factor Receptor  
; FILE REFERENCE: MBH00-876-K (400/021)  
; CURRENT APPLICATION NUMBER: US/09/685,664B  
; PRIOR FILING DATE: 2000-10-10  
; PRIOR APPLICATION NUMBER: US 60/005,974  
; PRIOR FILING DATE: 1995-10-26  
; PRIOR APPLICATION NUMBER: US 08/584,040  
; PRIOR FILING DATE: 1996-01-08  
; PRIOR APPLICATION NUMBER: US 09/371,772  
; PRIOR FILING DATE: 1999-08-10  
; NUMBER OF SEQ ID NOS: 8231  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 117  
; LENGTH: 17  
; TYPE: RNA  
; ORGANISM: Homo sapiens  
US-09-685-664B-117

Query Match 100.0%; Score 7; DB 6; Length 17;  
Best Local Similarity 100.0%; Pred. No. 7.1e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ttgtgata 7  
|||||||  
DB 14 TTTGATA 8

## RESULT 8

US-09-685-664B-118/C  
; Sequence 118, Application US/09685664B  
; GENERAL INFORMATION:  
; APPLICANT: Ribozyme Pharmaceuticals, Inc.  
; APPLICANT: Pavco, Pam  
; APPLICANT: MCSwigen, Jim  
; APPLICANT: Stinchcomb, Dan  
; APPLICANT: Escobedo, Jaime  
; TITLE OF INVENTION: Method and Reagent for Treatment of Diseases or Conditions Related to  
; TITLE OF INVENTION: Levels of Vascular Endothelial Growth Factor Receptor  
; FILE REFERENCE: MBH00-876-K (400/021)  
; CURRENT APPLICATION NUMBER: US/09/685,664B  
; PRIOR FILING DATE: 2000-10-10  
; PRIOR APPLICATION NUMBER: US 60/005,974  
; PRIOR FILING DATE: 1995-10-26  
; PRIOR APPLICATION NUMBER: US 08/584,040  
; PRIOR FILING DATE: 1996-01-08  
; PRIOR APPLICATION NUMBER: US 09/371,772  
; PRIOR FILING DATE: 1999-08-10  
; NUMBER OF SEQ ID NOS: 8231  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 118  
; LENGTH: 17  
; TYPE: RNA  
; ORGANISM: Homo sapiens  
US-09-685-664B-118

Query Match 100.0%; Score 7; DB 6; Length 17;  
Best Local Similarity 100.0%; Pred. No. 7.1e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0;  
Gaps 0;

QY 1 ttgtgata 7  
|||||||  
DB 12 TTTGATA 6

## RESULT 9

US-09-685-664B-882/C  
; Sequence 882, Application US/09685664B  
; GENERAL INFORMATION:  
; APPLICANT: Ribozyme Pharmaceuticals, Inc.  
; APPLICANT: Pavco, Pam  
; APPLICANT: MCSwigen, Jim  
; APPLICANT: Stinchcomb, Dan  
; APPLICANT: Escobedo, Jaime  
; TITLE OF INVENTION: Method and Reagent for Treatment of Diseases or Conditions Related to  
; TITLE OF INVENTION: Levels of Vascular Endothelial Growth Factor Receptor  
; FILE REFERENCE: MBH00-876-K (400/021)  
; CURRENT APPLICATION NUMBER: US/09/685,664B  
; PRIOR FILING DATE: 2000-10-10  
; PRIOR APPLICATION NUMBER: US 60/005,974  
; PRIOR FILING DATE: 1995-10-26  
; PRIOR APPLICATION NUMBER: US 08/584,040  
; PRIOR FILING DATE: 1996-01-08  
; PRIOR APPLICATION NUMBER: US 09/371,772  
; PRIOR FILING DATE: 1999-08-10  
; NUMBER OF SEQ ID NOS: 8231  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 882  
; LENGTH: 17  
; TYPE: RNA  
; ORGANISM: Homo sapiens  
US-09-685-664B-882

Query Match 100.0%; Score 7; DB 6; Length 17;  
Best Local Similarity 100.0%; Pred. No. 7.1e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ttgtgata 7  
|||||||  
DB 14 TTTGATA 8

## RESULT 10

US-09-685-664B-883/C  
; Sequence 883, Application US/09685664B  
; GENERAL INFORMATION:  
; APPLICANT: Ribozyme Pharmaceuticals, Inc.  
; APPLICANT: Pavco, Pam  
; APPLICANT: MCSwigen, Jim  
; APPLICANT: Stinchcomb, Dan  
; APPLICANT: Escobedo, Jaime  
; TITLE OF INVENTION: Method and Reagent for Treatment of Diseases or Conditions Related to  
; TITLE OF INVENTION: Levels of Vascular Endothelial Growth Factor Receptor  
; FILE REFERENCE: MBH00-876-K (400/021)  
; CURRENT APPLICATION NUMBER: US/09/685,664B  
; PRIOR FILING DATE: 2000-10-10  
; PRIOR APPLICATION NUMBER: US 60/005,974  
; PRIOR FILING DATE: 1995-10-26  
; PRIOR APPLICATION NUMBER: US 08/584,040  
; PRIOR FILING DATE: 1996-01-08  
; PRIOR APPLICATION NUMBER: US 09/371,772  
; PRIOR FILING DATE: 1999-08-10  
; NUMBER OF SEQ ID NOS: 8231  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 883  
; LENGTH: 17

TYPE: RNA  
ORGANISM: Homo sapiens  
US-09-685-664B-883

Query Match 100.0%; Score 7; DB 6; Length 17;  
Best Local Similarity 100.0%; Pred. No. 7.1e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttggata 7  
|||||||  
DB 12 TTTGATA 6

RESULT 11  
US-09-685-664B-2417/C  
Sequence 2417, Application US/09685664B  
GENERAL INFORMATION:  
APPLICANT: Ribozyne Pharmaceuticals, Inc.  
APPLICANT: Pavco, Pam  
APPLICANT: McSwiggen, Jim  
APPLICANT: Stinchcomb, Dan  
APPLICANT: Escobedo, Jaime  
TITLE OF INVENTION: Method and Reagent for Treatment of Diseases or Conditions Relate  
TITLE OF INVENTION: Levels of Vascular Endothelial Growth Factor Receptor  
FILE REFERENCE: MBH00-876-K (400/021)  
CURRENT APPLICATION NUMBER: US/09/685,664B  
PRIOR FILING DATE: 2000-10-10  
PRIOR APPLICATION NUMBER: US 60/005,974  
PRIOR FILING DATE: 1995-10-26  
PRIOR APPLICATION NUMBER: US 08/584,040  
PRIOR FILING DATE: 1996-01-08  
PRIOR APPLICATION NUMBER: US 09/371,772  
PRIOR FILING DATE: 1999-08-10  
NUMBER OF SEQ ID NOS: 8231  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 2417  
LENGTH: 17  
TYPE: RNA  
ORGANISM: Mus musculus  
US-09-685-664B-2417

Query Match 100.0%; Score 7; DB 6; Length 17;  
Best Local Similarity 100.0%; Pred. No. 7.1e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttggata 7  
|||||||  
DB 14 TTTGATA 8

RESULT 12  
US-09-685-664B-2418/C  
Sequence 2418, Application US/09685664B  
GENERAL INFORMATION:  
APPLICANT: Ribozyne Pharmaceuticals, Inc.  
APPLICANT: Pavco, Pam  
APPLICANT: McSwiggen, Jim  
APPLICANT: Stinchcomb, Dan  
APPLICANT: Escobedo, Jaime  
TITLE OF INVENTION: Method and Reagent for Treatment of Diseases or Conditions Relate  
TITLE OF INVENTION: Levels of Vascular Endothelial Growth Factor Receptor  
FILE REFERENCE: MBH00-876-K (400/021)  
CURRENT APPLICATION NUMBER: US/09/685,664B  
PRIOR FILING DATE: 2000-10-10  
PRIOR APPLICATION NUMBER: US 60/005,974  
PRIOR FILING DATE: 1995-10-26  
PRIOR APPLICATION NUMBER: US 08/584,040  
PRIOR FILING DATE: 1996-01-08  
PRIOR APPLICATION NUMBER: US 09/371,772  
PRIOR FILING DATE: 1999-08-10  
NUMBER OF SEQ ID NOS: 8231

SOFTWARE: PatentIn version 3.0  
SEQ ID NO 2418  
LENGTH: 17  
TYPE: RNA  
ORGANISM: Mus musculus  
US-09-685-664B-2418

Query Match 100.0%; Score 7; DB 6; Length 17;  
Best Local Similarity 100.0%; Pred. No. 7.1e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttggata 7  
|||||||  
DB 12 TTTGATA 6

RESULT 13  
US-09-685-664B-3108  
Sequence 3108, Application US/09685664B  
GENERAL INFORMATION:  
APPLICANT: Ribozyne Pharmaceuticals, Inc.  
APPLICANT: Pavco, Pam  
APPLICANT: McSwiggen, Jim  
APPLICANT: Stinchcomb, Dan  
APPLICANT: Escobedo, Jaime  
TITLE OF INVENTION: Method and Reagent for Treatment of Diseases or Conditions Rel  
TITLE OF INVENTION: Levels of Vascular Endothelial Growth Factor Receptor  
FILE REFERENCE: MBH00-876-K (400/021)  
CURRENT APPLICATION NUMBER: US/09/685,664B  
PRIOR FILING DATE: 2000-10-10  
PRIOR APPLICATION NUMBER: US 60/005,974  
PRIOR FILING DATE: 1995-10-26  
PRIOR APPLICATION NUMBER: US 08/584,040  
PRIOR FILING DATE: 1996-01-08  
PRIOR APPLICATION NUMBER: US 09/371,772  
PRIOR FILING DATE: 1999-08-10  
NUMBER OF SEQ ID NOS: 8231  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 3108  
LENGTH: 17  
TYPE: RNA  
ORGANISM: Mus musculus  
US-09-685-664B-3108

Query Match 100.0%; Score 7; DB 6; Length 17;  
Best Local Similarity 42.9%; Pred. No. 7.1e+04;  
Matches 3; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 ttggata 7  
:::||||  
DB 8 uuugaaua 14

RESULT 14  
US-09-685-664B-3109  
Sequence 3109, Application US/09685664B  
GENERAL INFORMATION:  
APPLICANT: Ribozyne Pharmaceuticals, Inc.  
APPLICANT: Pavco, Pam  
APPLICANT: McSwiggen, Jim  
APPLICANT: Stinchcomb, Dan  
APPLICANT: Escobedo, Jaime  
TITLE OF INVENTION: Method and Reagent for Treatment of Diseases or Conditions Rel  
TITLE OF INVENTION: Levels of Vascular Endothelial Growth Factor Receptor  
FILE REFERENCE: MBH00-876-K (400/021)  
CURRENT APPLICATION NUMBER: US/09/685,664B  
PRIOR FILING DATE: 2000-10-10  
PRIOR APPLICATION NUMBER: US 60/005,974  
PRIOR FILING DATE: 1995-10-26  
PRIOR APPLICATION NUMBER: US 08/584,040  
PRIOR FILING DATE: 1996-01-08

; PRIOR APPLICATION NUMBER: US 09/371,772  
 ; PRIOR FILING DATE: 1999-08-10  
 ; NUMBER OF SEQ ID NOS: 8231  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 3109  
 ; LENGTH: 17  
 ; TYPE: RNA  
 ; ORGANISM: Mus musculus  
 US-09-685-664B-3109

Query Match 100.0%; Score 7; DB 6; Length 17;  
 Best Local Similarity 42.9%; Pred. No. 7.1e+04;  
 Matches 3; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 ttgata 7  
 : : : | : | : |  
 Db 7 uuugaua 13

RESULT 15  
 US-09-685-664B-3110  
 ; Sequence 3110, Application US/09685664B  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ribozyne Pharmaceuticals, Inc.  
 ; APPLICANT: Pavco, Pam  
 ; APPLICANT: McSwigen, Jim  
 ; APPLICANT: Stinchcomb, Dan  
 ; APPLICANT: Escobedo, Jaime  
 ; TITLE OF INVENTION: Method and Reagent for Treatment of Diseases or Conditions Relate  
 ; FILE REFERENCE: MBHB00-876-K (400/021)  
 ; CURRENT APPLICATION NUMBER: US/09/685,664B  
 ; CURRENT FILING DATE: 2000-10-10  
 ; PRIOR APPLICATION NUMBER: US 60/005,974  
 ; PRIOR FILING DATE: 1995-10-26  
 ; PRIOR APPLICATION NUMBER: US 08/584,040  
 ; PRIOR FILING DATE: 1996-01-08  
 ; PRIOR APPLICATION NUMBER: US 09/371,772  
 ; PRIOR FILING DATE: 1999-08-10  
 ; NUMBER OF SEQ ID NOS: 8231  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 3110  
 ; LENGTH: 17  
 ; TYPE: RNA  
 ; ORGANISM: Mus musculus  
 US-09-685-664B-3110

Query Match 100.0%; Score 7; DB 6; Length 17;  
 Best Local Similarity 42.9%; Pred. No. 7.1e+04;  
 Matches 3; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 ttgata 7  
 : : : | : | : |  
 Db 3 uuugaua 9

Search completed: December 15, 2001, 05:58:20  
 Job time: 12263 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 05:42:57 ; Search time 7962.77 Seconds  
(without alignments) 48.423 Million cell updates/sec

Title: US-09-380-826a-6  
Perfect score: 22  
Sequence: 1 tfttgagannnnnnnttcgata-22

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 17159718 seqs, 8763200856 residues

Total number of hits satisfying chosen parameters: 34319436

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_NM\_Main:\*

```
1: /cgn2_6/ptodata/2/pna/PCYUS_COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US06_COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US07_COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US08_COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US081_COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US082_COMB.seq:*
7: /cgn2_6/ptodata/2/pna/US083_COMB.seq:*
8: /cgn2_6/ptodata/2/pna/US084_COMB.seq:*
9: /cgn2_6/ptodata/2/pna/US085_COMB.seq:*
10: /cgn2_6/ptodata/2/pna/US086_COMB.seq:*
11: /cgn2_6/ptodata/2/pna/US087_COMB.seq:*
12: /cgn2_6/ptodata/2/pna/US088_COMB.seq:*
13: /cgn2_6/ptodata/2/pna/US089_COMB.seq:*
14: /cgn2_6/ptodata/2/pna/US090_COMB.seq:*
15: /cgn2_6/ptodata/2/pna/US091_COMB.seq:*
16: /cgn2_6/ptodata/2/pna/US092_COMB.seq:*
17: /cgn2_6/ptodata/2/pna/US093_COMB.seq:*
18: /cgn2_6/ptodata/2/pna/US094_COMB.seq:*
19: /cgn2_6/ptodata/2/pna/US095_COMB.seq:*
20: /cgn2_6/ptodata/2/pna/US095A_COMB.seq:*
21: /cgn2_6/ptodata/2/pna/US095C_COMB.seq:*
22: /cgn2_6/ptodata/2/pna/US095D_COMB.seq:*
23: /cgn2_6/ptodata/2/pna/US096_COMB.seq:*
24: /cgn2_6/ptodata/2/pna/US096B_COMB.seq:*
25: /cgn2_6/ptodata/2/pna/US096C_COMB.seq:*
26: /cgn2_6/ptodata/2/pna/US096D_COMB.seq:*
27: /cgn2_6/ptodata/2/pna/US096E_COMB.seq:*
28: /cgn2_6/ptodata/2/pna/US097A_COMB.seq:*
29: /cgn2_6/ptodata/2/pna/US097B_COMB.seq:*
30: /cgn2_6/ptodata/2/pna/US097C_COMB.seq:*
31: /cgn2_6/ptodata/2/pna/US098_COMB.seq:*
32: /cgn2_6/ptodata/2/pna/US099_COMB.seq:*
33: /cgn2_6/ptodata/2/pna/US600_COMB.seq:*
34: /cgn2_6/ptodata/2/pna/US6001_COMB.seq:*
35: /cgn2_6/ptodata/2/pna/US6002_COMB.seq:*
36: /cgn2_6/ptodata/2/pna/US6003_COMB.seq:*
37: /cgn2_6/ptodata/2/pna/US6004_COMB.seq:*
38: /cgn2_6/ptodata/2/pna/US6005_COMB.seq:*
39: /cgn2_6/ptodata/2/pna/US6006_COMB.seq:*
40: /cgn2_6/ptodata/2/pna/US6007_COMB.seq:*
41: /cgn2_6/ptodata/2/pna/US6008_COMB.seq:*
42: /cgn2_6/ptodata/2/pna/US6009_COMB.seq:*
43: /cgn2_6/ptodata/2/pna/US6010_COMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	14	63.6	22	17	US-09-380-826a-6
2	14	63.6	22	17	US-09-380-826a-6
3	14	63.6	22	17	US-09-380-826a-6
4	14	63.6	147	18	US-09-497-191-767
5	14	63.6	243	11	US-08-725-863-2005
6	14	63.6	243	51	US-60-184-813-200
7	14	63.6	256	45	US-60-127-461-5765
8	14	63.6	259	12	US-08-812-505-2536
9	14	63.6	259	13	US-08-978-620-2536
10	14	63.6	259	34	US-60-013-529-2536
11	14	63.6	261	18	US-09-411-077-732
12	14	63.6	261	21	US-09-540-213-52141
13	14	63.6	261	43	US-60-103-748-618
14	14	63.6	317	30	US-09-770-255-141
15	14	63.6	322	28	US-09-705-926-17256
16	14	63.6	326	25	US-09-654-617-181276
17	14	63.6	326	27	US-09-684-016-181276
18	14	63.6	333	24	US-09-637-086A-25736
19	14	63.6	333	25	US-09-654-617-216399
20	14	63.6	333	27	US-09-684-016-216399
21	14	63.6	353	19	US-09-521-640-131370
22	14	63.6	365	14	US-60-211-750-49225
23	14	63.6	385	17	US-09-371-805-2913
24	14	63.6	385	17	US-09-371-805-2913
25	14	63.6	393	17	US-09-362-510-50353
26	14	63.6	393	17	US-09-362-510-50353
27	14	63.6	393	19	US-09-515-126-11383
28	14	63.6	396	22	US-09-577-409-11383
29	14	63.6	396	19	US-09-515-128-18844
30	14	63.6	398	17	US-09-362-510-50840
31	14	63.6	398	17	US-09-362-510A-50840
32	14	63.6	420	16	US-09-248-796-13484
33	14	63.6	420	42	US-60-096-409-13484
34	14	63.6	427	17	US-09-304-517A-90042
35	14	63.6	427	17	US-09-371-146A-90042
36	14	63.6	427	22	US-09-553-094-2550
37	14	63.6	427	25	US-09-654-617-292752
38	14	63.6	427	27	US-09-684-016-292752
39	14	63.6	427	46	US-60-130-180-1137
40	14	63.6	435	17	US-09-342-089-3899
41	14	63.6	435	17	US-09-342-069A-3899

```

c 42      14      63.6      456      53      US-60-207-458-38165      Sequence 38165, A
c 43      14      63.6      461      17      US-09-304-517A-85089      Sequence 85089, A
c 44      14      63.6      461      25      US-09-654-617-291806      Sequence 291806,
c 45      14      63.6      461      27      US-09-684-016-291806      Sequence 291806,

```

## ALIGNMENTS

```

RESULT 1
US-09-380-826A-2
; Sequence 2, Application US/09380826A
; GENERAL INFORMATION:
; APPLICANT: Chappel, Rod
; TITLE OF INVENTION: LEPTOSPIRA PATHOGENS
; FILE REFERENCE: DAVIE79.001APC
; CURRENT APPLICATION NUMBER: US/09/380,826A
; PRIOR FILING DATE: 1999-11-22
; PRIOR APPLICATION NUMBER: PCT/AU98/00145
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: AU P05494/97
; PRIOR FILING DATE: 1997-03-07
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Primer for L. fainei.
US-09-380-826A-2

```

```

Query Match      63.6%; Score 14; DB 17; Length 22;
Best Local Similarity 63.6%; Pred. No. 7.3e+02;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

```

QY 1 1 tgttgagannnnnnnttgata 22
    |||||
DB 1 1 tgttgatcacagaattgata 22

```

```

RESULT 2
US-09-380-826A-6
; Sequence 6, Application US/09380826A
; GENERAL INFORMATION:
; APPLICANT: Chappel, Rod
; TITLE OF INVENTION: LEPTOSPIRA PATHOGENS
; FILE REFERENCE: DAVIE79.001APC
; CURRENT APPLICATION NUMBER: US/09/380,826A
; PRIOR FILING DATE: 1999-11-22
; PRIOR APPLICATION NUMBER: PCT/AU98/00145
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: AU P05494/97
; PRIOR FILING DATE: 1997-03-07
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(22)
; OTHER INFORMATION: n = A,T,C or G
; OTHER INFORMATION: Primer for L. fainei.
US-09-380-826A-6

```

```

Query Match      63.6%; Score 14; DB 17; Length 22;
Best Local Similarity 100.0%; Pred. No. 7.3e+02;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 1 tgttgagannnnnnnttgata 22
    |||||
DB 1 1 tgttgagannnnnnnttgata 22

```

```

RESULT 3
US-09-380-826A-7
; Sequence 7, Application US/09380826A
; GENERAL INFORMATION:
; APPLICANT: Chappel, Rod
; TITLE OF INVENTION: LEPTOSPIRA PATHOGENS
; FILE REFERENCE: DAVIE79.001APC
; CURRENT APPLICATION NUMBER: US/09/380,826A
; PRIOR FILING DATE: 1999-11-22
; PRIOR APPLICATION NUMBER: PCT/AU98/00145
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: AU P05494/97
; PRIOR FILING DATE: 1997-03-07
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Primer for L. fainei.
US-09-380-826A-7

```

```

Query Match      63.6%; Score 14; DB 17; Length 22;
Best Local Similarity 63.6%; Pred. No. 7.3e+02;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

```

QY 1 1 tgttgagannnnnnnttgata 22
    |||||
DB 1 1 tgttgatcacagaattgata 22

```

```

RESULT 4
US-09-497-191-767/c
; Sequence 767, Application US/09497191
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai
; APPLICANT: CHEN, Xianfeng
; APPLICANT: SUBRAMANIAN, Gopalakrishnan
; APPLICANT: ZHENG, Liansheng
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS 84-89, 91-99
; FILE REFERENCE: 2750-694P
; CURRENT APPLICATION NUMBER: US/09/497,191
; CURRENT FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 920
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 767
; LENGTH: 147
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: any n or xaa = unknown
US-09-497-191-767

```

```

Query Match      63.6%; Score 14; DB 18; Length 147;
Best Local Similarity 63.6%; Pred. No. 1.1e+03;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

```

QY 1 1 tgttgagannnnnnnttgata 22
    |||||
DB 55 tgttgacttgacttggata 34

```

```

RESULT 5
US-08-725-863-2005
; Sequence 2005, Application US/08725863

```

```
; GENERAL INFORMATION:
; APPLICANT: Stuart, Susan G.
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Lane, John C.
; APPLICANT: Deleage, Angelo M.
; APPLICANT: Shable, James I.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
; TITLE OF INVENTION: HUMAN NEUTROPHILS
; NUMBER OF SEQUENCES: 3631
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/725, 863
; FILING DATE: Herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: CERRONE, MICHAEL C., Ph. D.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PD-0043 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 855-0555
; TELEFAX: (415) 845-4166
; INFORMATION FOR SEQ ID NO: 2005:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 243 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: 342263
; US-08-725-863-2005

Query Match          63.6%; Score 14; DB 11; Length 243;
Best Local Similarity 63.6%; Pred. No. 1.2e+03;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 tcttgagannnnnnnttgata 22
      |||||  |||||
Db 152 tcttgagatttctgttgata 173

RESULT 6
US-60-184-813-200
; Sequence 200, Application US/60184813
; GENERAL INFORMATION:
; APPLICANT: Hodgson, David M.
; APPLICANT: Lincoln, Stephen E.
; APPLICANT: Russo, Frank D.
; APPLICANT: Spillo, Peter A.
; APPLICANT: Banville, Steve C.
; APPLICANT: Bratcher, Shawn R.
; APPLICANT: Dufour, Gerard E.
; APPLICANT: Cohen, Howard J.
; APPLICANT: Rosen, Bruce
; APPLICANT: Shah, Purvi
; APPLICANT: Chai, Michael S.
; APPLICANT: Hillman, Jennifer L.
; TITLE OF INVENTION: NUCLEIC ACID SYNTHESIS AND MODIFICATION ENZYMES
; FILE REFERENCE: PT-0121 P
; CURRENT APPLICATION NUMBER: US/60/184, 813
; CURRENT FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 905
```

```
; SOFTWARE: PERL Program
; SEQ ID NO 200
; LENGTH: 243
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID NO: 1050812.1
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 130
; OTHER INFORMATION: a, t, c, g, or other
; US-60-184-813-200

Query Match          63.6%; Score 14; DB 51; Length 243;
Best Local Similarity 63.6%; Pred. No. 1.2e+03;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 tcttgagannnnnnnttgata 22
      |||||  |||||
Db 152 tcttgagatttctgttgata 173

RESULT 7
US-60-127-461-5765/C
; Sequence 5765, Application US/60127461
; GENERAL INFORMATION:
; APPLICANT: Chapman, Rowan
; APPLICANT: Thornton, Michael
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF CANDIDA ALBICANS
; FILE REFERENCE: PM-0007-2 P
; CURRENT APPLICATION NUMBER: US/60/127, 461
; CURRENT FILING DATE: 1999-04-01
; NUMBER OF SEQ ID NOS: 8247
; SOFTWARE: PERL Program
; SEQ ID NO 5765
; LENGTH: 256
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: AT808589484
; US-60-127-461-5765

Query Match          63.6%; Score 14; DB 45; Length 256;
Best Local Similarity 63.6%; Pred. No. 1.3e+03;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 tcttgagannnnnnnttgata 22
      |||||  |||||
Db 158 tcttgagatttctgttgata 137

RESULT 8
US-08-812-505-2536
; Sequence 2536, Application US/08812505
; GENERAL INFORMATION:
; APPLICANT: Gooding, Douglas H.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Akerblom, Ingrid E.
; APPLICANT: Deleage, Angelo M.
; APPLICANT: Naughton, Rebecca E.
; APPLICANT: Klingler, Tod M.
; APPLICANT: Kington, Michael S.
; APPLICANT: Hillman, Jennifer L.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
; NUMBER OF SEQUENCES: 3896
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
```

STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/812,505  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/013,529  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: LUTHER, BARBARA J.  
REGISTRATION NUMBER: 33954  
REFERENCE/DOCKET NUMBER: PD-0126P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 855-0555  
TELEFAX: (415) 852-0195  
INFORMATION FOR SEQ ID NO: 2536:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 259 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
IMMEDIATE SOURCE:  
CLONE: 865159  
US-08-812-505-2536

Query Match 63.6%; Score 14; DB 12; Length 259;  
Best Local Similarity 63.6%; Pred. No. 1.3e+03;  
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 ttgttgannnnnnnnttgata 22  
|||||  
DB 32 TGTTCGAAATTGTCATTGTGATA 53

RESULT 9  
US-08-978-620-2536  
Sequence 2536, Application US/08978620  
GENERAL INFORMATION:  
APPLICANT: Gooding, Douglas H.  
APPLICANT: Stuve, Laura L.  
APPLICANT: Stuart, Susan G.  
APPLICANT: Ito, Laura Y.  
APPLICANT: Akerbloom, Ingrid E.  
APPLICANT: Delegeane, Angelo M.  
APPLICANT: Naughton, Rebecca E.  
APPLICANT: Klingler, Tod M.  
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM  
TITLE OF INVENTION: HUMAN BRAIN  
NUMBER OF SEQUENCES: 7797  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 PORTER DRIVE  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/978,620  
FILING DATE: HERWITH

CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/032,151  
FILING DATE: DECEMBER 6, 1996  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/812,505  
FILING DATE: MARCH 7, 1997  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: LUTHER, BARBARA J.  
REGISTRATION NUMBER: 33954  
REFERENCE/DOCKET NUMBER: PD-0126-1 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 855-0555  
TELEFAX: (650) 852-0195  
INFORMATION FOR SEQ ID NO: 2536:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 259 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
IMMEDIATE SOURCE:  
CLONE: 865159  
US-08-978-620-2536

Query Match 63.6%; Score 14; DB 13; Length 259;  
Best Local Similarity 63.6%; Pred. No. 1.3e+03;  
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 ttgttgannnnnnnnttgata 22  
|||||  
DB 32 TGTTCGAAATTGTCATTGTGATA 53

RESULT 10  
US-60-013-529-2536  
Sequence 2536, Application US/60013529  
GENERAL INFORMATION:  
APPLICANT: Gooding, Douglas H.  
APPLICANT: Stuve, Laura L.  
APPLICANT: Stuart, Susan G.  
APPLICANT: Ito, Laura Y.  
APPLICANT: Akerbloom, Ingrid E.  
APPLICANT: Delegeane, Angelo M.  
APPLICANT: Naughton, Rebecca E.  
APPLICANT: Klingler, Tod M.  
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM  
TITLE OF INVENTION: ASTROCYTOMA  
NUMBER OF SEQUENCES: 3896  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 PORTER DRIVE  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/60/013,529  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: LUTHER, BARBARA J.  
REGISTRATION NUMBER: 33954  
REFERENCE/DOCKET NUMBER: PD-0126P  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 855-0555  
TELEFAX: (415) 852-0195  
INFORMATION FOR SEQ ID NO: 2536:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 259 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
CLONE: 865159  
US-60-013-529-2536

Query Match 63.6%; Score 14; DB 34; Length 259;  
Best Local Similarity 63.6%; Pred. No. 1.3e+03;  
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 tgttgannnnnnnttgata 22  
|||||  
Db 32 TGTGGATTTGTCATTGATA 53

RESULT 11  
US-09-411-077-732  
Sequence 732, Application US/09411077  
GENERAL INFORMATION:  
APPLICANT: Goralski, Thomas J.  
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM RAT LIVER  
FILE REFERENCE: P2-0096 US  
CURRENT APPLICATION NUMBER: US/09/411.077  
EARLIER FILING DATE: 1999-10-04  
EARLIER APPLICATION NUMBER: 60/103.748  
EARLIER FILING DATE: October 9, 1998  
EARLIER APPLICATION NUMBER: 60/111.900  
EARLIER FILING DATE: December 10, 1998  
NUMBER OF SEQ ID NOS: 1651  
SOFTWARE: PERL Program  
SEQ ID NO 732  
LENGTH: 261  
TYPE: DNA  
ORGANISM: Rattus norvegicus  
FEATURE: -  
OTHER INFORMATION: 701257277H1  
US-09-411-077-732

Query Match 63.6%; Score 14; DB 18; Length 261;  
Best Local Similarity 63.6%; Pred. No. 1.3e+03;  
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 tgttgannnnnnnttgata 22  
|||||  
Db 202 tgttgaaaatcctcttgata 223

RESULT 12  
US-09-540-213-52141  
Sequence 52141, Application US/09540213  
GENERAL INFORMATION:  
APPLICANT: Seilhamer, Jeffrey J.  
APPLICANT: Deleage, Angelo M.  
APPLICANT: Stuart, Susan G.  
APPLICANT: Stuve, Laura L.  
APPLICANT: Mullahy, Sara J.  
APPLICANT: Naughton, Rebecca E.  
TITLE OF INVENTION: POLYNUCLEOTIDES OF LIVER AND PANCREAS  
FILE REFERENCE: PD-1031 CIP  
CURRENT APPLICATION NUMBER: US/09/540.213  
CURRENT FILING DATE: 2000-03-31  
Prior application data removed - consult PALM or file wrapper  
NUMBER OF SEQ ID NOS: 53844

SEQ ID NO 52141  
LENGTH: 261  
TYPE: DNA  
ORGANISM: Rattus norvegicus  
FEATURE:  
NAME/KEY: misc-feature  
OTHER INFORMATION: Incyte ID No: rat00178536  
US-09-540-213-52141

Query Match 63.6%; Score 14; DB 21; Length 261;  
Best Local Similarity 63.6%; Pred. No. 1.3e+03;  
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 tgttgannnnnnnttgata 22  
|||||  
Db 202 tgttgaaaatcctcttgata 223

RESULT 13  
US-60-103-748-618  
Sequence 618, Application US/60103748  
GENERAL INFORMATION:  
APPLICANT: Stuve, Laura L.  
APPLICANT: Goralski, Thomas J.  
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM TREATED RAT LIVER  
FILE REFERENCE: P2-0096 P  
CURRENT APPLICATION NUMBER: US/60/103.748  
CURRENT FILING DATE: 1998-10-09  
NUMBER OF SEQ ID NOS: 1537  
SOFTWARE: PERL Program  
SEQ ID NO 618  
LENGTH: 261  
TYPE: DNA  
ORGANISM: Rattus norvegicus  
FEATURE: -  
OTHER INFORMATION: 701257277H1  
US-60-103-748-618

Query Match 63.6%; Score 14; DB 43; Length 261;  
Best Local Similarity 63.6%; Pred. No. 1.3e+03;  
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 tgttgannnnnnnttgata 22  
|||||  
Db 202 tgttgaaaatcctcttgata 223

RESULT 14  
US-09-770-255-141  
Sequence 141, Application US/09770255  
GENERAL INFORMATION:  
APPLICANT: Gorlach, Jorn  
APPLICANT: An, Yong-Qiang  
APPLICANT: Hamilton, Carol M.  
APPLICANT: Price, Jennifer L.  
APPLICANT: Raines, Tracy M.  
APPLICANT: Yu, Yang  
APPLICANT: Rameaka, Joshua G.  
APPLICANT: Page, Amy  
APPLICANT: Mathew, Abraham V.  
APPLICANT: Ledford, Brooke L.  
APPLICANT: Woessner, Jeffrey P.  
APPLICANT: Haas, William David  
APPLICANT: Garcia, Carlos A.  
APPLICANT: Krickler, Maja  
APPLICANT: Slader, Ted  
APPLICANT: Davis, Keith R.  
APPLICANT: Allen, Keith  
APPLICANT: Hoffman, Neil  
APPLICANT: Hurban, Patrick

```

; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; FILE OF INVENTION: thaliana
; FILE REFERENCE: 2030US (PARA-019PRV)
; CURRENT APPLICATION NUMBER: US/09/770,255
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,504
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 141
; LENGTH: 317
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(317)
; OTHER INFORMATION: n = A,T,C or G
US-09-770-255-141

```

```

Query Match          63.6%; Score 14; DB 30; Length 317;
Best Local Similarity 63.6%; Pred. No. 1.3e+03;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

```

OY 1 tgttgannnnnnnnttgata 22
    |||||
Db 167 tgttgatgtatggttgata 188

```

```

RESULT 15
US-09-705-926-17256/C
; Sequence 17256, Application US/09705926
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jindong
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: ANNOTATED PLANT GENES
; FILE REFERENCE: 38-21(15481)A
; CURRENT APPLICATION NUMBER: US/09/705,926
; CURRENT FILING DATE: 2000-11-06
; NUMBER OF SEQ ID NOS: 21634
; SEQ ID NO 17256
; LENGTH: 322
; TYPE: DNA
; ORGANISM: Lycopersicon esculentum
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(322)
; OTHER INFORMATION: unsure at all n locations
US-09-705-926-17256

```

```

Query Match          63.6%; Score 14; DB 28; Length 322;
Best Local Similarity 63.6%; Pred. No. 1.3e+03;
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

```

OY 1 tgttgannnnnnnnttgata 22
    |||||
Db 112 TGTTCGAGGGTTCTTTGATA 91

```

Search completed: December 15, 2001, 05:42:59  
Job time: 15824 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 05:58:20 ; Search time 474.83 Seconds  
(without alignments)  
66.629 Million cell updates/sec

Title: US-09-380-826a-6

Perfect score: 1 lgltgannnnnnnttgata 22

Sequence:

Scoring table: IDENTITY\_NUC

Searched: Gapop 10.0 , Gapext 1.0

Total number of hits satisfying chosen parameters: 1946350

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Pending\_Patents\_NA\_New:\*  
1: /cgn2\_6/ptodata/1/pna/PCR\_NEW\_COMB.seq:\*  
2: /cgn2\_6/ptodata/1/pna/US06\_NEW\_COMB.seq:\*  
3: /cgn2\_6/ptodata/1/pna/US07\_NEW\_COMB.seq:\*  
4: /cgn2\_6/ptodata/1/pna/US08\_NEW\_COMB.seq:\*  
5: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:\*  
6: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq1:\*  
7: /cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:\*  
8: /cgn2\_6/ptodata/1/pna/US60\_NEW\_COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	14	63.6	427	6	US-09-985-678-90042
C 2	14	63.6	461	6	US-09-985-678-85089
C 3	14	63.6	857	6	US-09-922-279A-3008
C 4	14	63.6	1287	1	PCT-US01-15674A-50
C 5	14	63.6	1501	6	US-09-815-264-12871
C 6	14	63.6	1501	6	US-09-815-264-14471
C 7	14	63.6	1501	6	US-09-815-264-50077
C 8	14	63.6	1501	6	US-09-815-264-50328
C 9	14	63.6	6926	6	US-09-815-264-60943
C 10	14	63.6	7316	6	US-09-815-264-69885
C 11	14	63.6	7673	6	US-09-815-264-74657
C 12	14	63.6	41742	6	US-09-815-264-59220
C 13	14	63.6	94612	6	US-09-815-264-79259
C 14	14	59.1	226	6	US-09-815-242-1444
C 15	14	59.1	245	6	US-09-894-949A-151011
C 16	14	59.1	300	6	US-09-894-949A-1500
C 17	14	59.1	320	6	US-09-815-242-3155
C 18	14	59.1	378	6	US-09-815-242-2111
C 19	14	59.1	378	6	US-09-985-678-231502
C 20	14	59.1	388	6	US-09-388-906A-23742
C 21	14	59.1	425	6	US-09-933-524A-18588
C 22	14	59.1	424	6	US-09-991-423-606
C 23	14	59.1	443	6	US-09-933-524A-110244
C 24	14	59.1	443	6	US-09-815-264-29506
C 25	14	59.1	464	6	US-09-815-242-1944

26	13	59.1	471	6	US-09-933-524A-109421	Sequence 109421,
C 27	13	59.1	501	6	US-09-985-678-220277	Sequence 220277,
C 28	13	59.1	709	6	US-09-815-264-65443	Sequence 65443, A
C 29	13	59.1	845	6	US-09-922-279A-1268	Sequence 1268, Ap
C 30	13	59.1	918	6	US-09-815-242-4840	Sequence 4840, Ap
C 31	13	59.1	1092	6	US-09-922-279A-3103	Sequence 3103, Ap
C 32	13	59.1	1263	6	US-09-815-242-4274	Sequence 4274, Ap
C 33	13	59.1	1266	6	US-09-815-242-6090	Sequence 8090, Ap
C 34	13	59.1	1266	6	US-09-815-242-8747	Sequence 8747, Ap
C 35	13	59.1	1415	6	US-09-898-888A-14137	Sequence 14137, A
C 36	13	59.1	1481	6	US-09-815-264-106283	Sequence 106283,
C 37	13	59.1	1501	6	US-09-815-264-7540	Sequence 7540, Ap
C 38	13	59.1	1537	6	US-09-815-264-25196	Sequence 25196, A
C 39	13	59.1	1537	6	US-09-815-264-106280	Sequence 106280,
C 40	13	59.1	3209	6	US-09-897-516-970	Sequence 970, App
C 41	13	59.1	3498	6	US-09-815-242-6647	Sequence 6647, Ap
C 42	13	59.1	3627	6	US-09-815-264-91108	Sequence 91108, A
C 43	13	59.1	7391	6	US-09-815-264-58271	Sequence 58271, A
C 44	13	59.1	9854	6	US-09-815-264-66587	Sequence 66587, A
C 45	13	59.1	10040	6	US-09-815-264-78579	Sequence 78579, A

#### ALIGNMENTS

RESULT 1  
US-09-985-678-90042/C  
Sequence 90042, Application US/09985678  
GENERAL INFORMATION:  
APPLICANT: Cheikh, Nordine  
TITLE OF INVENTION: Annotated Plant Genes  
FILE REFERENCE: 16517.255/38-21(15097)F  
CURRENT APPLICATION NUMBER: US/09/985,678  
CURRENT FILING DATE: 2001-11-05  
PRIOR APPLICATION NUMBER: US 09/304,517  
PRIOR FILING DATE: 1999-05-06  
NUMBER OF SEQ ID NOS: 295529  
SEQ ID NO 90042  
LENGTH: 427  
TYPE: DNA  
ORGANISM: Zea mays  
US-09-985-678-90042

Query Match 63.6%; Score 14; DB 6; Length 427;  
Best Local Similarity 63.6%; Pred. NO. 2.2e+02;  
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 lgltgannnnnnnttgata 22  
|||||  
Db 262 TGTTCGATGCACAGTTTGATA 241

RESULT 2  
US-09-985-678-85089/C  
Sequence 85089, Application US/09985678  
GENERAL INFORMATION:  
APPLICANT: Cheikh, Nordine  
TITLE OF INVENTION: Annotated Plant Genes  
FILE REFERENCE: 16517.255/38-21(15097)F  
CURRENT APPLICATION NUMBER: US/09/985,678  
CURRENT FILING DATE: 2001-11-05  
PRIOR APPLICATION NUMBER: US 09/304,517  
PRIOR FILING DATE: 1999-05-06  
NUMBER OF SEQ ID NOS: 295529  
SEQ ID NO 85089  
LENGTH: 461  
TYPE: DNA  
ORGANISM: Zea mays  
US-09-985-678-85089

Query Match 63.6%; Score 14; DB 6; Length 461;  
Best Local Similarity 63.6%; Pred. No. 2.2e+02;  
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 tgttgagannnnnnnttgata 22  
|||||  
DB 300 TGTGTGAGAGACACGTTTGATA 279

## RESULT 3

US-09-922-279A-3008  
; Sequence 3008, Application US/09922279A  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Yuanhua T.  
; APPLICANT: Tillinghast, John  
; APPLICANT: Slinku, Ankura  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Drmanac, Radoje T.  
; TITLE OF INVENTION: Novel Contigs Obtained  
; FILE REFERENCE: 785  
; CURRENT APPLICATION NUMBER: US/09/922,279A  
; CURRENT FILING DATE: 2001-08-03  
; PRIOR APPLICATION NUMBER: 09/491,404  
; PRIOR FILING DATE: 2000-01-25  
; NUMBER OF SEQ ID NOS: 3796  
; SOFTWARE: PL\_SP\_genes Version 1.0  
; SEQ ID NO 3008  
; LENGTH: 857  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: sig\_peptide  
; LOCATION: (514)..(649)  
; OTHER INFORMATION: this location contains the signal peptide sequence.  
; OTHER INFORMATION: MGAFVTRVALPCSFAMCACLIGPASALCSLCMKNFQFFOYLDS, Run with SignalP  
; NAME/KEY: misc\_feature  
; LOCATION: (290)..(213)  
; OTHER INFORMATION: similar to g12231098 in the genepept database release 114,  
; OTHER INFORMATION: Run with FASTRY 3.3c00, default parameters  
US-09-922-279A-3008

Query Match 63.6%; Score 14; DB 6; Length 857;  
Best Local Similarity 63.6%; Pred. No. 2.4e+02;  
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 tgttgagannnnnnnttgata 22  
|||||  
DB 685 tgttgagaggaagtcgttgata 706

## RESULT 4

PCT-US01-15674A-50  
; Sequence 50, Application PC/TUS0115674A  
; GENERAL INFORMATION:  
; APPLICANT: JOAN, KNOLL H  
; APPLICANT: ROGAN, PETER K  
; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATING  
; FILE REFERENCE: 30307  
; CURRENT APPLICATION NUMBER: PCT/US01/15674A  
; CURRENT FILING DATE: 2001-12-07  
; NUMBER OF SEQ ID NOS: 613  
; SOFTWARE: Patent version 3.1  
; SEQ ID NO 50  
; LENGTH: 1287  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: repeat\_region  
; LOCATION: (1)..(1287)  
; OTHER INFORMATION: hsmar1

## PCT-US01-15674A-50

Query Match 63.6%; Score 14; DB 1; Length 1287;  
Best Local Similarity 63.6%; Pred. No. 2.5e+02;  
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 tgttgagannnnnnnttgata 22  
|||||  
DB 36 tgttgagattgcgcttgata 57

## RESULT 5

US-09-815-264-12871  
; Sequence 12871, Application US/09815264  
; GENERAL INFORMATION:  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Dotson, Stanton B.  
; APPLICANT: Koshi, Jeffrey M.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Liu, Jingdong  
; APPLICANT: McIninch, James  
; APPLICANT: Wu, Wei  
; TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof  
; FILE REFERENCE: 38-21(51237)G  
; CURRENT APPLICATION NUMBER: US/09/815,264  
; CURRENT FILING DATE: 2001-03-23  
; PRIOR APPLICATION NUMBER: US 09/620,392  
; PRIOR FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: US 09/702,134  
; PRIOR FILING DATE: 2000-10-31  
; NUMBER OF SEQ ID NOS: 109669  
; SEQ ID NO 12871  
; LENGTH: 1501  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
US-09-815-264-12871

Query Match 63.6%; Score 14; DB 6; Length 1501;  
Best Local Similarity 63.6%; Pred. No. 2.6e+02;  
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 tgttgagannnnnnnttgata 22  
|||||  
DB 671 tgttgagaggaatttgata 692

## RESULT 6

US-09-815-264-14471  
; Sequence 14471, Application US/09815264  
; GENERAL INFORMATION:  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Dotson, Stanton B.  
; APPLICANT: Koshi, Jeffrey M.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Liu, Jingdong  
; APPLICANT: McIninch, James  
; APPLICANT: Wu, Wei  
; TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof  
; FILE REFERENCE: 38-21(51237)G  
; CURRENT APPLICATION NUMBER: US/09/815,264  
; CURRENT FILING DATE: 2001-03-23  
; PRIOR APPLICATION NUMBER: US 09/620,392  
; PRIOR FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: US 09/702,134  
; PRIOR FILING DATE: 2000-10-31  
; NUMBER OF SEQ ID NOS: 109669  
; SEQ ID NO 14471  
; LENGTH: 1501  
; TYPE: DNA



ORGANISM: Oryza sativa  
US-09-815-264-14471

Query Match 63.6%; Score 14; DB 6; Length 1501;  
Best Local Similarity 63.6%; Pred. No. 2.6e+02;  
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 tgttgannnnnnnttgata 22  
||||| |||||||  
Db 671 tcttgacaggaatttgata 692

RESULT 7  
US-09-815-264-50077/C  
Sequence 50077, Application US/09815264  
GENERAL INFORMATION:  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Cao, Yongwei  
APPLICANT: Dotson, Stanton B.  
APPLICANT: Koshi, Jeffrey M.  
APPLICANT: Kovalic, David K.  
APPLICANT: Liu, Jindong  
APPLICANT: McIninch, James  
APPLICANT: Wu, Wei  
TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof  
FILE REFERENCE: 38-21(51237)G  
CURRENT APPLICATION NUMBER: US/09/815,264  
CURRENT FILING DATE: 2001-03-23  
PRIOR APPLICATION NUMBER: US 09/620,392  
PRIOR FILING DATE: 2000-07-19  
PRIOR APPLICATION NUMBER: US 09/702,134  
PRIOR FILING DATE: 2000-10-31  
NUMBER OF SEQ ID NOS: 109669  
SEQ ID NO 50077  
LENGTH: 1501  
TYPE: DNA  
ORGANISM: Oryza sativa  
US-09-815-264-50077

Query Match 63.6%; Score 14; DB 6; Length 1501;  
Best Local Similarity 63.6%; Pred. No. 2.6e+02;  
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 tgttgannnnnnnttgata 22  
||||| |||||||  
Db 1311 TGTGATATTATCTTGATA 1290

RESULT 8  
US-09-815-264-50328  
Sequence 50328, Application US/09815264  
GENERAL INFORMATION:  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Cao, Yongwei  
APPLICANT: Dotson, Stanton B.  
APPLICANT: Koshi, Jeffrey M.  
APPLICANT: Kovalic, David K.  
APPLICANT: Liu, Jindong  
APPLICANT: McIninch, James  
APPLICANT: Wu, Wei  
TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof  
FILE REFERENCE: 38-21(51237)G  
CURRENT APPLICATION NUMBER: US/09/815,264  
CURRENT FILING DATE: 2001-03-23  
PRIOR APPLICATION NUMBER: US 09/620,392  
PRIOR FILING DATE: 2000-07-19  
PRIOR APPLICATION NUMBER: US 09/702,134  
PRIOR FILING DATE: 2000-10-31  
NUMBER OF SEQ ID NOS: 109669  
SEQ ID NO 50328  
LENGTH: 1501

TYPE: DNA  
ORGANISM: Oryza sativa  
US-09-815-264-50328

Query Match 63.6%; Score 14; DB 6; Length 1501;  
Best Local Similarity 63.6%; Pred. No. 2.6e+02;  
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 tgttgannnnnnnttgata 22  
||||| |||||||  
Db 124 tcttgatattatccttgata 145

RESULT 9  
US-09-815-264-60943  
Sequence 60943, Application US/09815264  
GENERAL INFORMATION:  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Cao, Yongwei  
APPLICANT: Dotson, Stanton B.  
APPLICANT: Koshi, Jeffrey M.  
APPLICANT: Kovalic, David K.  
APPLICANT: Liu, Jindong  
APPLICANT: McIninch, James  
APPLICANT: Wu, Wei  
TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof  
FILE REFERENCE: 38-21(51237)G  
CURRENT APPLICATION NUMBER: US/09/815,264  
CURRENT FILING DATE: 2001-03-23  
PRIOR APPLICATION NUMBER: US 09/620,392  
PRIOR FILING DATE: 2000-07-19  
PRIOR APPLICATION NUMBER: US 09/702,134  
PRIOR FILING DATE: 2000-10-31  
NUMBER OF SEQ ID NOS: 109669  
SEQ ID NO 60943  
LENGTH: 6926  
TYPE: DNA  
ORGANISM: Oryza sativa  
US-09-815-264-60943

Query Match 63.6%; Score 14; DB 6; Length 6926;  
Best Local Similarity 63.6%; Pred. No. 3.2e+02;  
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 tgttgannnnnnnttgata 22  
||||| |||||||  
Db 4746 tcttgacaggaatttgata 4767

RESULT 10  
US-09-815-264-69885/C  
Sequence 69885, Application US/09815264  
GENERAL INFORMATION:  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Cao, Yongwei  
APPLICANT: Dotson, Stanton B.  
APPLICANT: Koshi, Jeffrey M.  
APPLICANT: Kovalic, David K.  
APPLICANT: Liu, Jindong  
APPLICANT: McIninch, James  
APPLICANT: Wu, Wei  
TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof  
FILE REFERENCE: 38-21(51237)G  
CURRENT APPLICATION NUMBER: US/09/815,264  
CURRENT FILING DATE: 2001-03-23  
PRIOR APPLICATION NUMBER: US 09/620,392  
PRIOR FILING DATE: 2000-07-19  
PRIOR APPLICATION NUMBER: US 09/702,134  
PRIOR FILING DATE: 2000-10-31  
NUMBER OF SEQ ID NOS: 109669  
SEQ ID NO 69885

LENGTH: 7316  
TYPE: DNA  
ORGANISM: Oryza sativa  
US-09-815-264-69885

Query Match 63.6%; Score 14; DB 6; Length 7316;  
Best Local Similarity 63.6%; Pred. No. 3.2e+02;  
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 tgttgannnnnnnttgata 22  
|||||  
Db 3151 TGTGGATTGTTGATTGATA 3130

RESULT 11  
US-09-815-264-74657/C  
Sequence 74657, Application US/09815264  
GENERAL INFORMATION:  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Cao, Yongwei  
APPLICANT: Dotson, Stanton B.  
APPLICANT: Koshi, Jeffrey M.  
APPLICANT: Kovalic, David K.  
APPLICANT: Liu, Jingdong  
APPLICANT: McIninch, James  
APPLICANT: Wu, Wei  
TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof  
FILE REFERENCE: 38-21(51237)G  
CURRENT APPLICATION NUMBER: US/09/815,264  
CURRENT FILING DATE: 2001-03-23  
PRIOR APPLICATION NUMBER: US 09/620,392  
PRIOR FILING DATE: 2000-07-19  
PRIOR APPLICATION NUMBER: US 09/702,134  
PRIOR FILING DATE: 2000-10-31  
NUMBER OF SEQ ID NOS: 109669  
SEQ ID NO 74657  
LENGTH: 7673  
TYPE: DNA  
ORGANISM: Oryza sativa  
US-09-815-264-74657

Query Match 63.6%; Score 14; DB 6; Length 7673;  
Best Local Similarity 63.6%; Pred. No. 3.2e+02;  
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 tgttgannnnnnnttgata 22  
|||||  
Db 3586 TGTGGACATATATTGATA 3565

RESULT 12  
US-09-815-264-59220  
Sequence 59220, Application US/09815264  
GENERAL INFORMATION:  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Cao, Yongwei  
APPLICANT: Dotson, Stanton B.  
APPLICANT: Koshi, Jeffrey M.  
APPLICANT: Kovalic, David K.  
APPLICANT: Liu, Jingdong  
APPLICANT: McIninch, James  
APPLICANT: Wu, Wei  
TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof  
FILE REFERENCE: 38-21(51237)G  
CURRENT APPLICATION NUMBER: US/09/815,264  
CURRENT FILING DATE: 2001-03-23  
PRIOR APPLICATION NUMBER: US 09/620,392  
PRIOR FILING DATE: 2000-07-19  
PRIOR APPLICATION NUMBER: US 09/702,134  
PRIOR FILING DATE: 2000-10-31  
NUMBER OF SEQ ID NOS: 109669

SEQ ID NO 59220  
LENGTH: 41742  
TYPE: DNA  
ORGANISM: Oryza sativa  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)..(41742)  
OTHER INFORMATION: unsure at all n locations  
US-09-815-264-59220

Query Match 63.6%; Score 14; DB 6; Length 41742;  
Best Local Similarity 63.6%; Pred. No. 4.1e+02;  
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 tgttgannnnnnnttgata 22  
|||||  
Db 2402 tgttgaacaggaatttgata 2423

RESULT 13  
US-09-815-264-79259  
Sequence 79259, Application US/09815264  
GENERAL INFORMATION:  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Cao, Yongwei  
APPLICANT: Dotson, Stanton B.  
APPLICANT: Koshi, Jeffrey M.  
APPLICANT: Kovalic, David K.  
APPLICANT: Liu, Jingdong  
APPLICANT: McIninch, James  
APPLICANT: Wu, Wei  
TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof  
FILE REFERENCE: 38-21(51237)G  
CURRENT APPLICATION NUMBER: US/09/815,264  
CURRENT FILING DATE: 2001-03-23  
PRIOR APPLICATION NUMBER: US 09/620,392  
PRIOR FILING DATE: 2000-07-19  
PRIOR APPLICATION NUMBER: US 09/702,134  
PRIOR FILING DATE: 2000-10-31  
NUMBER OF SEQ ID NOS: 109669  
SEQ ID NO 79259  
LENGTH: 94612  
TYPE: DNA  
ORGANISM: Oryza sativa  
US-09-815-264-79259

Query Match 63.6%; Score 14; DB 6; Length 94612;  
Best Local Similarity 63.6%; Pred. No. 4.5e+02;  
Matches 14; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 tgttgannnnnnnttgata 22  
|||||  
Db 14796 tgttgatattatccttgata 14817

RESULT 14  
US-09-815-242-1444  
Sequence 1444, Application US/09815242  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Karl L.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: Identification of Essential Genes in  
FILE REFERENCE: ELITRA.011A  
CURRENT APPLICATION NUMBER: US/09/815,242

```

: CURRENT FILING DATE: 2001-03-21
: PRIOR APPLICATION NUMBER: 60/191,078
: PRIOR FILING DATE: 2000-03-21
: PRIOR APPLICATION NUMBER: 60/206,848
: PRIOR FILING DATE: 2000-05-23
: PRIOR APPLICATION NUMBER: 60/207,727
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: 60/242,578
: PRIOR FILING DATE: 2000-10-23
: PRIOR APPLICATION NUMBER: 60/253,625
: PRIOR FILING DATE: 2000-11-27
: PRIOR APPLICATION NUMBER: 60/257,931
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: 60/269,308
: PRIOR FILING DATE: 2001-02-16
: NUMBER OF SEQ ID NOS: 14110
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1444
: LENGTH: 226
: TYPE: DNA
: ORGANISM: Staphylococcus aureus
US-09-815-242-1444

```

```

Query Match          59.1%; Score 13; DB 6; Length 226;
Best Local Similarity 61.9%; Pred. No. 6.9e+02;
Matches 13; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 2 gtggannnnnnnnttgata 22
    |||||
Db 122 gtggatgtagctcttgata 142

```

```

RESULT 15
US-09-985-678-151011/C
: Sequence 151011, Application US/09985678
: GENERAL INFORMATION:
: APPLICANT: Cheikh, Nordine
: TITLE OF INVENTION: Annotated Plant Genes
: FILE REFERENCE: 16517.255/38-21(15087)F
: CURRENT APPLICATION NUMBER: US/09/985,678
: CURRENT FILING DATE: 2001-11-05
: PRIOR APPLICATION NUMBER: US 09/304,517
: PRIOR FILING DATE: 1999-05-06
: NUMBER OF SEQ ID NOS: 295529
: SEQ ID NO 151011
: LENGTH: 245
: TYPE: DNA
: ORGANISM: Glycine max
US-09-985-678-151011

```

```

Query Match          59.1%; Score 13; DB 6; Length 245;
Best Local Similarity 61.9%; Pred. No. 6.9e+02;
Matches 13; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 2 gtggannnnnnnnttgata 22
    |||||
Db 165 GTTGCATTGTATCATTTGATA 145

```

Search completed: December 15, 2001, 05:58:24  
Job time: 12267 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 05:42:59 ; Search time 7962.77 Seconds  
(without alignments)  
48.423 Million cell updates/sec

Title: US-09-380-826a-7

Perfect score: 22

Sequence: 1 ttttgatcaccaagatttgata 22

Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 17159718 seqs, 8763200856 residues

Total number of hits satisfying chosen parameters: 34319436

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending\_Patents\_NA\_Main:\*

1: /cgn2\_6/ptodata/2/pna/US0611\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/pna/US06\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/pna/US07\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/pna/US080\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/pna/US081\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/pna/US082\_COMB.seq:\*  
7: /cgn2\_6/ptodata/2/pna/US083\_COMB.seq:\*  
8: /cgn2\_6/ptodata/2/pna/US084\_COMB.seq:\*  
9: /cgn2\_6/ptodata/2/pna/US085\_COMB.seq:\*  
10: /cgn2\_6/ptodata/2/pna/US086\_COMB.seq:\*  
11: /cgn2\_6/ptodata/2/pna/US087\_COMB.seq:\*  
12: /cgn2\_6/ptodata/2/pna/US088\_COMB.seq:\*  
13: /cgn2\_6/ptodata/2/pna/US089\_COMB.seq:\*  
14: /cgn2\_6/ptodata/2/pna/US090\_COMB.seq:\*  
15: /cgn2\_6/ptodata/2/pna/US091\_COMB.seq:\*  
16: /cgn2\_6/ptodata/2/pna/US093\_COMB.seq:\*  
17: /cgn2\_6/ptodata/2/pna/US093\_COMB.seq:\*  
18: /cgn2\_6/ptodata/2/pna/US094\_COMB.seq:\*  
19: /cgn2\_6/ptodata/2/pna/US095A\_COMB.seq:\*  
20: /cgn2\_6/ptodata/2/pna/US095B\_COMB.seq:\*  
21: /cgn2\_6/ptodata/2/pna/US095C\_COMB.seq:\*  
22: /cgn2\_6/ptodata/2/pna/US095D\_COMB.seq:\*  
23: /cgn2\_6/ptodata/2/pna/US096A\_COMB.seq:\*  
24: /cgn2\_6/ptodata/2/pna/US096B\_COMB.seq:\*  
25: /cgn2\_6/ptodata/2/pna/US096C\_COMB.seq:\*  
26: /cgn2\_6/ptodata/2/pna/US096D\_COMB.seq:\*  
27: /cgn2\_6/ptodata/2/pna/US096E\_COMB.seq:\*  
28: /cgn2\_6/ptodata/2/pna/US097A\_COMB.seq:\*  
29: /cgn2\_6/ptodata/2/pna/US097B\_COMB.seq:\*  
30: /cgn2\_6/ptodata/2/pna/US097C\_COMB.seq:\*  
31: /cgn2\_6/ptodata/2/pna/US098\_COMB.seq:\*  
32: /cgn2\_6/ptodata/2/pna/US099\_COMB.seq:\*  
33: /cgn2\_6/ptodata/2/pna/US6000\_COMB.seq:\*  
34: /cgn2\_6/ptodata/2/pna/US6001\_COMB.seq:\*  
35: /cgn2\_6/ptodata/2/pna/US6002\_COMB.seq:\*  
36: /cgn2\_6/ptodata/2/pna/US6003\_COMB.seq:\*  
37: /cgn2\_6/ptodata/2/pna/US6004\_COMB.seq:\*  
38: /cgn2\_6/ptodata/2/pna/US6005\_COMB.seq:\*  
39: /cgn2\_6/ptodata/2/pna/US6006\_COMB.seq:\*  
40: /cgn2\_6/ptodata/2/pna/US6007\_COMB.seq:\*  
41: /cgn2\_6/ptodata/2/pna/US6008\_COMB.seq:\*  
42: /cgn2\_6/ptodata/2/pna/US6009\_COMB.seq:\*  
43: /cgn2\_6/ptodata/2/pna/US6010\_COMB.seq:\*

44: /cgn2\_6/ptodata/2/pna/US6011\_COMB.seq:\*  
45: /cgn2\_6/ptodata/2/pna/US6012\_COMB.seq:\*  
46: /cgn2\_6/ptodata/2/pna/US6013\_COMB.seq:\*  
47: /cgn2\_6/ptodata/2/pna/US6014\_COMB.seq:\*  
48: /cgn2\_6/ptodata/2/pna/US6015\_COMB.seq:\*  
49: /cgn2\_6/ptodata/2/pna/US6016\_COMB.seq:\*  
50: /cgn2\_6/ptodata/2/pna/US6017\_COMB.seq:\*  
51: /cgn2\_6/ptodata/2/pna/US6018\_COMB.seq:\*  
52: /cgn2\_6/ptodata/2/pna/US6019\_COMB.seq:\*  
53: /cgn2\_6/ptodata/2/pna/US6020\_COMB.seq:\*  
54: /cgn2\_6/ptodata/2/pna/US6021\_COMB.seq:\*  
55: /cgn2\_6/ptodata/2/pna/US6022\_COMB.seq:\*  
56: /cgn2\_6/ptodata/2/pna/US6023\_COMB.seq:\*  
57: /cgn2\_6/ptodata/2/pna/US6024\_COMB.seq:\*  
58: /cgn2\_6/ptodata/2/pna/US6025\_COMB.seq:\*  
59: /cgn2\_6/ptodata/2/pna/US6026\_COMB.seq:\*  
60: /cgn2\_6/ptodata/2/pna/US6027\_COMB.seq:\*  
61: /cgn2\_6/ptodata/2/pna/US6028\_COMB.seq:\*  
62: /cgn2\_6/ptodata/2/pna/US6029\_COMB.seq:\*  
63: /cgn2\_6/ptodata/2/pna/US6030\_COMB.seq:\*  
64: /cgn2\_6/ptodata/2/pna/US6031\_COMB.seq:\*  
65: /cgn2\_6/ptodata/2/pna/US6032\_COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	22	100.0	22	17	US-09-380-826a-7	Sequence 2, Appl
2	22	100.0	22	17	US-09-380-826a-1	Sequence 1, Appl
3	18.8	85.5	420	46	US-60-132-861-11944	Sequence 11944, A
4	18.8	85.5	530	25	US-09-649-165A-7425	Sequence 7425, Ap
5	18.8	85.5	581	25	US-09-654-617-175385	Sequence 175385,
6	18.8	85.5	581	27	US-09-684-016-175385	Sequence 175385,
7	18.8	85.5	649	22	US-09-565-309A-20998	Sequence 20998, A
8	18.8	85.5	1069	22	US-09-565-309A-57118	Sequence 57118, A
9	18.8	85.5	1072	22	US-09-565-309A-49885	Sequence 49885, A
10	18.8	85.5	1128	25	US-09-654-617-117747	Sequence 117747,
11	18.8	85.5	1429	46	US-60-138-103-11564	Sequence 11564, A
12	18.8	85.5	5874	18	US-09-404-550-5289	Sequence 5289, Ap
13	18.8	85.5	92624	20	US-09-534-859-352	Sequence 352, App
14	18.8	85.5	92624	31	US-09-803-736-352	Sequence 352, App
15	18.4	83.6	2463	52	US-60-143-993-148	Sequence 148, App
16	18.4	83.6	2463	52	US-60-191-637-27704	Sequence 27704, A
17	18.4	83.6	2463	52	US-60-191-637-22313	Sequence 22313, A
18	18.4	83.6	2463	52	US-60-167-217-19072	Sequence 19072, A
19	18.4	83.6	2463	52	US-60-167-217-19008	Sequence 19008, A
20	18.4	83.6	2463	52	US-60-191-637-15063	Sequence 15063, A
21	18.4	83.6	2463	52	US-60-191-637-18351	Sequence 18351, A
22	18.4	83.6	2463	52	US-60-191-637-14552	Sequence 14552, A
23	18.4	83.6	2463	52	US-60-167-217-18359	Sequence 18359, A
24	18.4	83.6	2463	52	US-09-528-237A-1225	Sequence 1225, Ap
25	18.4	83.6	2463	52	US-09-654-617-180482	Sequence 180482,
26	18.4	83.6	2463	52	US-09-684-016-180482	Sequence 180482,
27	18.4	83.6	2463	52	US-09-521-640-103292	Sequence 103292,
28	18.4	83.6	2463	52	US-09-521-640-103292	Sequence 103292,
29	18.4	83.6	2463	52	US-09-521-640-103292	Sequence 103292,
30	18.4	83.6	2463	52	US-09-521-640-103292	Sequence 103292,
31	18.4	83.6	2463	52	US-09-521-640-103292	Sequence 103292,
32	18.4	83.6	2463	52	US-09-521-640-103292	Sequence 103292,
33	18.4	83.6	2463	52	US-09-521-640-103292	Sequence 103292,
34	18.4	83.6	2463	52	US-09-521-640-103292	Sequence 103292,
35	18.4	83.6	2463	52	US-09-521-640-103292	Sequence 103292,
36	18.4	83.6	2463	52	US-09-521-640-103292	Sequence 103292,
37	18.4	83.6	2463	52	US-09-521-640-103292	Sequence 103292,
38	18.4	83.6	2463	52	US-09-521-640-103292	Sequence 103292,
39	18.4	83.6	2463	52	US-09-521-640-103292	Sequence 103292,
40	18.4	83.6	2463	52	US-09-521-640-103292	Sequence 103292,
41	18.4	83.6	2463	52	US-09-521-640-103292	Sequence 103292,
42	18.4	83.6	2463	52	US-09-521-640-103292	Sequence 103292,
43	18.4	83.6	2463	52	US-09-521-640-103292	Sequence 103292,

42	17.2	78.2	249	24	US-09-638-258-520	Sequence 520, App
43	17.2	78.2	249	32	US-09-924-035-520	Sequence 520, App
44	17.2	78.2	278	25	US-09-654-617-370318	Sequence 370318,
45	17.2	78.2	278	27	US-09-684-016-370318	Sequence 370318,

## ALIGNMENTS

```
RESULT 1
US-09-380-826a-7
; Sequence 2, Application US/09380826A
; GENERAL INFORMATION:
; APPLICANT: Chappel, Rod
; TITLE OF INVENTION: LEPTOSPIRA PATHOGENS
; FILE REFERENCE: DAVIE79.001APC
; CURRENT APPLICATION NUMBER: US/09/380,826A
; CURRENT FILING DATE: 1999-11-22
; PRIOR APPLICATION NUMBER: PCT/AU98/00145
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: AU P05494/97
; PRIOR FILING DATE: 1997-03-07
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Primer for L. fainei.
US-09-380-826a-2
```

```
Query Match          100.0%; Score 22; DB 17; Length 22;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 tgttgatcacagaatttgata 22
    |||
Db 1 tgttgatcacagaatttgata 22
```

```
RESULT 2
US-09-380-826a-7
; Sequence 7, Application US/09380826A
; GENERAL INFORMATION:
; APPLICANT: Chappel, Rod
; TITLE OF INVENTION: LEPTOSPIRA PATHOGENS
; FILE REFERENCE: DAVIE79.001APC
; CURRENT APPLICATION NUMBER: US/09/380,826A
; CURRENT FILING DATE: 1999-11-22
; PRIOR APPLICATION NUMBER: PCT/AU98/00145
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: AU P05494/97
; PRIOR FILING DATE: 1997-03-07
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Primer for L. fainei.
US-09-380-826a-7
```

```
Query Match          100.0%; Score 22; DB 17; Length 22;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 tgttgatcacagaatttgata 22
    |||
Db 1 tgttgatcacagaatttgata 22
```

```
RESULT 3
US-09-380-826a-1
; Sequence 1, Application US/09380826A
; GENERAL INFORMATION:
; APPLICANT: Chappel, Rod
; TITLE OF INVENTION: LEPTOSPIRA PATHOGENS
; FILE REFERENCE: DAVIE79.001APC
; CURRENT APPLICATION NUMBER: US/09/380,826A
; CURRENT FILING DATE: 1999-11-22
; PRIOR APPLICATION NUMBER: PCT/AU98/00145
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: AU P05494/97
; PRIOR FILING DATE: 1997-03-07
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1477
; TYPE: DNA
; ORGANISM: Leptospira fainei
US-09-380-826a-1
```

```
Query Match          100.0%; Score 22; DB 17; Length 1477;
Best Local Similarity 100.0%; Pred. No. 3,3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 tgttgatcacagaatttgata 22
    |||
Db 154 tgttgatcacagaatttgata 175
```

```
RESULT 4
US-60-132-861-11944
; Sequence 11944, Application US/60132861
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Timberlake, William E.
; TITLE OF INVENTION: ASPERGILLUS NIDULANS GENOME SEQUENCE AND USES THEREOF
; FILE REFERENCE: 38-21(15485)C
; CURRENT APPLICATION NUMBER: US/60/132,861
; CURRENT FILING DATE: 1999-05-07
; NUMBER OF SEQ ID NOS: 13800
; SEQ ID NO 11944
; LENGTH: 420
; TYPE: DNA
; ORGANISM: Aspergillus nidulans
US-60-132-861-11944
```

```
Query Match          85.5%; Score 18.8; DB 46; Length 420;
Best Local Similarity 90.9%; Pred. No. 1,1e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 1 tgttgatcacagaatttgata 22
    |||
Db 234 tgttgatcacagaatttgata 255
```

```
RESULT 5
US-09-649-165a-7425/c
; Sequence 7425, Application US/09649165A
; GENERAL INFORMATION:
; APPLICANT: Guegler, Karl J.
; APPLICANT: Doyle, Martin
; APPLICANT: Momiyama, Monika
; APPLICANT: Wang, Xiaohao
; TITLE OF INVENTION: DETECTING EXPRESSION OF ARABIDOPSIS THALIANA GENES
; FILE REFERENCE: PA-0019 US
; CURRENT APPLICATION NUMBER: US/09/649,165A
; CURRENT FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 60/150,517
```

PRIOR FILING DATE: 1999-08-24  
; NUMBER OF SEQ ID NOS: 7753  
; SOFTWARE: PERL Program  
; SEQ ID NO 7425  
; LENGTH: 530  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc.feature  
; OTHER INFORMATION: Incycle ID No: 701676626  
; NAME/KEY: unsure  
; LOCATION: 6  
; OTHER INFORMATION: a, t, c, g, dr other  
US-09-649-165A-7425

Query Match 85.5%; Score 18.8; DB 25; Length 530;  
Best Local Similarity 90.9%; Pred. No. 1.1e+02;  
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tgttgatcacagaatttgata 22  
||||| |||||||||  
Db 254 tcttggttcacaaagatttcaca 233

RESULT 6  
US-09-654-617-175385  
; Sequence 175385, Application US/09654617  
; GENERAL INFORMATION:  
; APPLICANT: Kovallc, David K.  
; APPLICANT: Liu, Jingdong  
; TITLE OF INVENTION: Annotated Plant Genes  
; FILE REFERENCE: 38-21(15097)D  
; CURRENT APPLICATION NUMBER: US/09/654,617  
; CURRENT FILING DATE: 2000-09-05  
; NUMBER OF SEQ ID NOS: 463173  
; SEQ ID NO 175385  
; LENGTH: 581  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-654-617-175385

Query Match 85.5%; Score 18.8; DB 25; Length 581;  
Best Local Similarity 90.9%; Pred. No. 1.1e+02;  
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tgttgatcacagaatttgata 22  
||||| |||||||||  
Db 420 tgttgatcacagaatttgata 441

RESULT 7  
US-09-684-016-175385  
; Sequence 175385, Application US/09684016  
; GENERAL INFORMATION:  
; APPLICANT: Kovallc, David K.  
; APPLICANT: Liu, Jingdong  
; TITLE OF INVENTION: Annotated Plant Genes  
; FILE REFERENCE: 38-21(15097)D  
; CURRENT APPLICATION NUMBER: US/09/684,016  
; CURRENT FILING DATE: 2000-10-10  
; PRIOR APPLICATION NUMBER: US 09/654,617  
; PRIOR FILING DATE: 2000-09-05  
; NUMBER OF SEQ ID NOS: 463173  
; SEQ ID NO 175385  
; LENGTH: 581  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-684-016-175385

Query Match 85.5%; Score 18.8; DB 27; Length 581;

Best Local Similarity 90.9%; Pred. No. 1.1e+02;  
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tgttgatcacagaatttgata 22  
||||| |||||||||  
Db 420 tgttgatcacagaatttgata 441

RESULT 8  
US-09-565-309A-20908  
; Sequence 20908, Application US/09565309A  
; GENERAL INFORMATION:  
; APPLICANT: ALEXANDROV, Nikolai  
; APPLICANT: BROVER, Vyacheslav  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
; FILE REFERENCE: 2750-0853p  
; CURRENT APPLICATION NUMBER: US/09/565,309A  
; CURRENT FILING DATE: 2000-05-05  
; NUMBER OF SEQ ID NOS: 68449  
; SEQ ID NO 20908  
; LENGTH: 649  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (1)..(649)  
; OTHER INFORMATION: any n = a, g, c, t, unknown, or other  
; NAME/KEY: (1)..(649)  
; LOCATION: (1)..(649)  
; OTHER INFORMATION: 27496;110925 (Clone Number:Unique Sequence Identifier)  
US-09-565-309A-20908

Query Match 85.5%; Score 18.8; DB 22; Length 649;  
Best Local Similarity 90.9%; Pred. No. 1.1e+02;  
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tgttgatcacagaatttgata 22  
||||| |||||||||  
Db 265 tgttgatcacagaatttgata 286

RESULT 9  
US-09-565-309A-57118  
; Sequence 57118, Application US/09565309A  
; GENERAL INFORMATION:  
; APPLICANT: ALEXANDROV, Nikolai  
; APPLICANT: BROVER, Vyacheslav  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
; FILE REFERENCE: 2750-0853p  
; CURRENT APPLICATION NUMBER: US/09/565,309A  
; CURRENT FILING DATE: 2000-05-05  
; NUMBER OF SEQ ID NOS: 68449  
; SEQ ID NO 57118  
; LENGTH: 1069  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (1)..(1069)  
; OTHER INFORMATION: any n = a, g, c, t, unknown, or other  
; NAME/KEY: misc.feature  
; LOCATION: (1)..(1069)  
; OTHER INFORMATION: 27496 ; OVERLAP (Clone Number : OVERLAP)  
US-09-565-309A-57118

Query Match 85.5%; Score 18.8; DB 22; Length 1069;  
Best Local Similarity 90.9%; Pred. No. 1.2e+02;  
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tttgtgacacaagatttgata 22  
||||| |||||||||  
Db 690 tttgtgacacaagatttgaca 711

## RESULT 10

US-09-565-309A-49885  
; Sequence 49885, Application US/09565309A  
; GENERAL INFORMATION:  
; APPLICANT: ALEXANDROV, Nikolai  
; APPLICANT: BROVER, Vyacheslav  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; FILE REFERENCE: 2750-0853P  
; CURRENT APPLICATION NUMBER: US/09/565,309A  
; CURRENT FILING DATE: 2000-05-05  
; NUMBER OF SEQ ID NOS: 68449  
; SEQ ID NO 49885  
; LENGTH: 1072  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(1072)  
; OTHER INFORMATION: any n = a, g, c, t, unknown, or other  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(1072)  
; OTHER INFORMATION: 27496 : 5TAG CONSENSUS (Clone Number:5tag\_consensus)  
US-09-565-309A-49885

Query Match 85.5%; Score 18.8; DB 22; Length 1072;  
Best Local Similarity 90.9%; Pred. No. 1.2e+02;  
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tttgtgacacaagatttgata 22  
||||| |||||||||  
Db 691 tttgtgacacaagatttgaca 712

## RESULT 11

US-09-654-617-117747  
; Sequence 117747, Application US/09654617  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Liu, Jingdong  
; TITLE OF INVENTION: Annotated Plant Genes  
; FILE REFERENCE: 38-21(15097)D  
; CURRENT APPLICATION NUMBER: US/09/654,617  
; CURRENT FILING DATE: 2000-09-05  
; NUMBER OF SEQ ID NOS: 463173  
; SEQ ID NO 117747  
; LENGTH: 1128  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
; OTHER INFORMATION: unsure at all n locations  
US-09-654-617-117747

Query Match 85.5%; Score 18.8; DB 25; Length 1128;  
Best Local Similarity 90.9%; Pred. No. 1.2e+02;  
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tttgtgacacaagatttgata 22  
||||| |||||||||  
Db 719 tttgtgacacaagatttgaca 740

RESULT 12  
US-09-684-016-117747  
; Sequence 117747, Application US/09684016  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.

; APPLICANT: Liu, Jingdong  
; TITLE OF INVENTION: Annotated Plant Genes  
; FILE REFERENCE: 38-21(15097)D  
; CURRENT APPLICATION NUMBER: US/09/684,016  
; CURRENT FILING DATE: 2000-10-10  
; PRIOR APPLICATION NUMBER: US 09/654,617  
; PRIOR FILING DATE: 2000-09-05  
; NUMBER OF SEQ ID NOS: 463173  
; SEQ ID NO 117747  
; LENGTH: 1128  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(1128)  
; OTHER INFORMATION: unsure at all n locations  
US-09-684-016-117747

Query Match 85.5%; Score 18.8; DB 27; Length 1128;  
Best Local Similarity 90.9%; Pred. No. 1.2e+02;  
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tttgtgacacaagatttgata 22  
||||| |||||||||  
Db 719 tttgtgacacaagatttgaca 740

## RESULT 13

US-60-138-103-11564/C  
; Sequence 11564, Application US/60138103  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Timberlake, William E.  
; TITLE OF INVENTION: ASPERGILLUS NIDULANS GENOME SEQUENCE AND USES THEREOF  
; FILE REFERENCE: 38-10(15485)D  
; CURRENT APPLICATION NUMBER: US/60/138,103  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 28006  
; SEQ ID NO 11564  
; LENGTH: 1429  
; TYPE: DNA  
; ORGANISM: Aspergillus nidulans  
US-60-138-103-11564

Query Match 85.5%; Score 18.8; DB 46; Length 1429;  
Best Local Similarity 90.9%; Pred. No. 1.2e+02;  
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tttgtgacacaagatttgata 22  
||||| |||||||||  
Db 1206 tttgtgacacaagatttgaca 1185

RESULT 14  
US-09-404-520-5289  
; Sequence 5289, Application US/09404520  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Ghodssi, Azita  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: McIninch, James  
; APPLICANT: Timberlake, William E.  
; APPLICANT: Yu, Jaehyuk  
; TITLE OF INVENTION: EMERICELLA NIDULANS GENOME SEQUENCE AND USES THEREOF  
; FILE REFERENCE: 38-10(15498)A  
; CURRENT APPLICATION NUMBER: US/09/404,520  
; CURRENT FILING DATE: 1999-09-23  
; NUMBER OF SEQ ID NOS: 44345  
; SEQ ID NO 5289  
; LENGTH: 5874  
; TYPE: DNA



: ORGANISM: Aspergillus nidulans  
US-09-404-520-5289

Query Match 85.5%; Score 18.8; DB 18; Length 5874;  
Best Local Similarity 90.9%; Pred. No. 1.4e+02;  
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tgttgatcacagaatttgata 22  
||||| ||||| ||||| |||||  
Db 248 tgttgatcacagaatttgata 269

RESULT 15

US-09-534-859-352  
: Sequence 352, Application US/09534859  
: GENERAL INFORMATION:  
: APPLICANT: Bush, David F.  
: APPLICANT: Last, Robert L.  
: APPLICANT: Levin, Irena M.  
: APPLICANT: Norris, Susan R.  
: APPLICANT: Parnell, Laurence D.  
: APPLICANT: Rounsley, Steven D.  
: APPLICANT: Wiegand, Roger C.  
: TITLE OF INVENTION: PLANT POLYMORPHIC MARKERS AND USES THEREOF  
: FILE REFERENCE: 38-10(15493)B  
: CURRENT APPLICATION NUMBER: US/09/534,859  
: CURRENT FILING DATE: 2000-03-29  
: NUMBER OF SEQ ID NOS: 1127  
: SEQ ID NO 352  
: LENGTH: 92624  
: TYPE: DNA  
: ORGANISM: Arabidopsis thaliana  
US-09-534-859-352

Query Match 85.5%; Score 18.8; DB 20; Length 92624;  
Best Local Similarity 90.9%; Pred. No. 2e+02;  
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tgttgatcacagaatttgata 22  
||||| ||||| ||||| |||||  
Db 35807 tgttgatcacagaatttgata 35828

Search completed: December 15, 2001, 05:43:01  
Job time: 15826 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 05:58:24 ; Search time 474.83 seconds  
(without alignments)  
66.629 Million cell updates/sec

Title: US-09-380-826a-7

Perfect score: 22

Sequence: 1 tgttgatcaccaagattgata 22

Scoring table: IDENTITY\_NUC

Gapop 10.0, Gapext 1.0

Searched: 973175 seqs, 719035916 residues

Total number of hits satisfying chosen parameters: 1946350

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_NA\_New:\*  
1: /cg2\_6/ptodata/1/pna/PCr\_NEW\_COMB.seq:\*  
2: /cg2\_6/ptodata/1/pna/US06\_NEW\_COMB.seq:\*  
3: /cg2\_6/ptodata/1/pna/US07\_NEW\_COMB.seq:\*  
4: /cg2\_6/ptodata/1/pna/US08\_NEW\_COMB.seq:\*  
5: /cg2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:\*  
6: /cg2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:\*  
7: /cg2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:\*  
8: /cg2\_6/ptodata/1/pna/US60\_NEW\_COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	18.8	85.5	1069	US-09-620-394B-4668	Sequence 4668, Ap
2	17.2	78.2	249	US-09-924-035A-520	Sequence 520, App
3	17.2	78.2	622	US-09-985-678-211824	Sequence 211824, Ap
4	17.2	78.2	2038	US-09-922-279A-2773	Sequence 2773, Ap
5	17.2	78.2	46891	US-09-815-264-82805	Sequence 82805, A
6	16.8	76.4	267	US-09-982-402-1036	Sequence 1036, Ap
7	16.8	76.4	551	US-09-849-526A-27340	Sequence 27340, A
8	16.8	76.4	849	US-09-815-264-26466	Sequence 26466, A
9	16.8	76.4	1046	US-09-815-264-63319	Sequence 63319, A
10	16.8	76.4	1246	US-09-815-264-91672	Sequence 91672, A
11	16.8	76.4	1501	US-09-815-264-23140	Sequence 23140, A
12	16.8	76.4	1501	US-09-815-264-46882	Sequence 46882, A
13	16.8	76.4	1501	US-09-815-264-53250	Sequence 53250, A
14	16.8	76.4	6940	US-09-815-264-76889	Sequence 76889, A
15	16.8	76.4	12081	US-09-815-264-81045	Sequence 81045, A
16	16.8	76.4	12198	US-09-815-264-76653	Sequence 76653, A
17	16.8	76.4	41269	US-09-815-264-59808	Sequence 59808, A
18	16.8	76.4	50929	US-09-815-264-61033	Sequence 61033, A
19	16.8	76.4	67216	US-09-815-264-71944	Sequence 71944, A
20	16.4	74.5	360	US-09-388-906A-17536	Sequence 17536, A
21	16.4	74.5	4861	US-09-815-264-109466	Sequence 109466, A
22	16.2	73.6	278	US-09-985-678-279880	Sequence 279880, A
23	16.2	73.6	376	US-09-912-293-18469	Sequence 18469, A
24	16.2	73.6	377	US-09-845-487A-238	Sequence 238, App
25	16.2	73.6	419	US-09-933-524A-2094	Sequence 2094, Ap

26	16.2	73.6	461	US-09-849-526A-12135	Sequence 12135, A
27	16.2	73.6	845	US-09-922-279A-1268	Sequence 1268, Ap
28	16.2	73.6	1251	US-09-898-888A-9009	Sequence 9009, Ap
29	16.2	73.6	1457	US-09-815-264-60838	Sequence 60838, A
30	16.2	73.6	1501	US-09-815-264-39775	Sequence 39775, A
31	16.2	73.6	1501	US-09-815-264-39776	Sequence 39776, A
32	16.2	73.6	1501	US-09-815-264-43531	Sequence 43531, A
33	16.2	73.6	3603	US-09-815-264-80803	Sequence 80803, A
34	16.2	73.6	6237	US-09-815-264-72243	Sequence 72243, A
35	16.2	73.6	7856	US-09-815-264-86401	Sequence 86401, A
36	16.2	73.6	9723	US-09-815-264-75693	Sequence 75693, A
37	16.2	73.6	10839	US-09-815-264-78946	Sequence 78946, A
38	16.2	73.6	26601	US-09-815-264-74669	Sequence 74669, A
39	16.2	73.6	33005	US-09-815-264-73015	Sequence 73015, A
40	16.2	73.6	36227	US-09-815-264-68619	Sequence 68619, A
41	16	72.7	610	US-09-985-678-213635	Sequence 213635, A
42	16	72.7	1381	US-09-815-264-108378	Sequence 108378, A
43	16	72.7	37658	US-09-815-264-74259	Sequence 74259, A
44	16	72.7	39365	US-09-815-264-79204	Sequence 79204, A
45	15.8	71.8	270	US-09-985-678-42368	Sequence 42368, A

## ALIGNMENTS

RESULT 1  
US-09-620-394B-4668  
Sequence 4668, Application US/09620394B  
GENERAL INFORMATION:  
APPLICANT: ALEXANDROV, Nikolai  
TITLE OF INVENTION: DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDE  
TITLE OF INVENTION: Thereby  
FILE REFERENCE: 2750-1067P  
CURRENT APPLICATION NUMBER: US/09/620.394B  
CURRENT FILING DATE: 2000-07-21  
NUMBER OF SEQ ID NOS: 9131  
SEQ ID NO 4668  
LENGTH: 1069  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: 1..1069  
OTHER INFORMATION: any n = a, g, c, t, unknown, or other  
NAME/KEY: misc.feature  
LOCATION: 1..1069  
OTHER INFORMATION: Ceres Seq. ID 1393173  
US-09-620-394B-4668  
Query Match 85.5%; Score 18.8; DB 6; Length 1069;  
Best Local Similarity 90.9%; Pred. No. 16;  
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
OY 1 tgttgatcaccaagattgata 22  
DB 690 tgttgatcaccaagattgata 711  
RESULT 2  
US-09-924-035A-520  
Sequence 520, Application US/09924035A  
GENERAL INFORMATION:  
APPLICANT: Grilach, Jrm  
TITLE OF INVENTION: EXPRESSED SEQUENCES OF ARABIDOPSIS  
TITLE OF INVENTION: thaliana  
FILE REFERENCE: 2011US  
CURRENT APPLICATION NUMBER: US/09/924.035A  
CURRENT FILING DATE: 2000-08-11  
PRIOR APPLICATION NUMBER: US 60/148.784  
PRIOR FILING DATE: 1999-08-13  
NUMBER OF SEQ ID NOS: 900

```

; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 520
; LENGTH: 249
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(249)
; OTHER INFORMATION: n = A,T,C or G
US-09-924-035A-520

```

```

Query Match      78.2%; Score 17.2; DB 5; Length 249;
Best Local Similarity 86.4%; Pred. No. 81;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1 tgttgatcacagaatttgata 22
||||| ||||||||| |||
DB 54 tcttgtaacacaagaatttgta 75

```

```

RESULT 3
US-09-985-678-211824
; Sequence 211824, Application US/09985678
; GENERAL INFORMATION:
; APPLICANT: Cheikh, Norline
; APPLICANT: Liu, Jingdong
; TITLE OF INVENTION: Annotated Plant Genes
; FILE REFERENCE: 16517.255/38-21(15097)F
; CURRENT APPLICATION NUMBER: US/09/985,678
; CURRENT FILING DATE: 2001-11-05
; PRIOR APPLICATION NUMBER: US 09/304,517
; PRIOR FILING DATE: 1999-05-06
; NUMBER OF SEQ ID NOS: 295529
; SEQ ID NO 211824
; LENGTH: 622
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)...(622)
; OTHER INFORMATION: unsure at all n locations
US-09-985-678-211824

```

```

Query Match      78.2%; Score 17.2; DB 6; Length 622;
Best Local Similarity 86.4%; Pred. No. 87;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1 tgttgatcacagaatttgata 22
||||| ||||||||| |||||||
DB 520 tgatgatactcaagcttgata 541

```

```

RESULT 4
US-09-922-279A-2773
; Sequence 2773, Application US/09922279A
; GENERAL INFORMATION:
; APPLICANT: Tang, Yuanhua T.
; APPLICANT: Tillinghast, John
; APPLICANT: Sinks, Ankura
; APPLICANT: Liu, Chenghua
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Contigs Obtained
; FILE REFERENCE: 785
; CURRENT APPLICATION NUMBER: US/09/922,279A
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/491,404
; PRIOR FILING DATE: 2000-01-25
; NUMBER OF SEQ ID NOS: 3796
; SOFTWARE: PL_SP_genes Version 1.0
; SEQ ID NO 2773

```

```

; LENGTH: 2038
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: (1091)...(956)
; OTHER INFORMATION: this location contains the signal peptide sequence
; OTHER INFORMATION: MLYLLFPGVSYLRSLFGRPIGPIGTSDFLLFLFSNLDSPPLS, Run with Signal
; NAME/KEY: misc_feature
; LOCATION: (605)...(1902)
; OTHER INFORMATION: similar to g15931821 in the genepept database release 114,
; OTHER INFORMATION: Run with FASTX 3.3t00, default parameters
US-09-922-279A-2773

```

```

Query Match      78.2%; Score 17.2; DB 6; Length 2038;
Best Local Similarity 86.4%; Pred. No. 96;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1 tgttgatcacagaatttgata 22
||||| ||||||||| |||
DB 1929 tcttgagcacagaattttaa 1950

```

```

RESULT 5
US-09-815-264-82805
; Sequence 82805, Application US/09815264
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Cao, Yongwei
; APPLICANT: Dotson, Stanton B.
; APPLICANT: Koshi, Jeffrey M.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: McIninch, James
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51237)G
; CURRENT APPLICATION NUMBER: US/09/815,264
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 09/620,392
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/702,134
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 109669
; SEQ ID NO 82805
; LENGTH: 46891
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-815-264-82805

```

```

Query Match      78.2%; Score 17.2; DB 6; Length 46891;
Best Local Similarity 86.4%; Pred. No. 1,3e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1 tgttgatcacagaatttgata 22
||||| ||||||||| |||||||
DB 45218 tcttgacacacaagaatttgata 45239

```

```

RESULT 6
US-09-982-402-1036
; Sequence 1036, Application US/09982402
; GENERAL INFORMATION:
; APPLICANT: [list inventors here]
; TITLE OF INVENTION: [list title here]
; FILE REFERENCE: [list docket ID here]
; CURRENT APPLICATION NUMBER: US/09/982,402
; CURRENT FILING DATE: 2001-10-17
; NUMBER OF SEQ ID NOS: 3553
; SOFTWARE: PERL Program
; SEQ ID NO 1036

```

LENGTH: 267  
TYPE: DNA  
ORGANISM: Zea mays  
FEATURE:  
NAME/KEY: misc\_feature  
OTHER INFORMATION: Incyte ID No: 700196524H1  
US-09-982-402-1036

Query Match 76.4%; Score 16.8; DB 6; Length 267;  
Best Local Similarity 90.0%; Pred. No. 1.3e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 gtggatcacaaagttagat 21  
||||| ||||| |||  
Db 180 gtggatgcacaaagttagat 199

RESULT 7  
US-09-849-526A-27340/C  
Sequence 27340, Application US/09849526A  
GENERAL INFORMATION:  
APPLICANT: Ruff, Thomas G.  
TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With  
FILE REFERENCE: 16517,250138-21(51930)B)  
CURRENT APPLICATION NUMBER: US/09/849,526A  
PRIOR FILING DATE: 2001-05-07  
PRIOR APPLICATION NUMBER: US 60/202,214  
PRIOR FILING DATE: 2000-05-08  
PRIOR APPLICATION NUMBER: US 09/654,617  
PRIOR FILING DATE: 2000-09-05  
PRIOR APPLICATION NUMBER: US 09/684,016  
PRIOR FILING DATE: 2000-10-10  
PRIOR APPLICATION NUMBER: US 09/816,660  
PRIOR FILING DATE: 2001-03-26  
NUMBER OF SEQ ID NOS: 30131  
SEQ ID NO 27340  
LENGTH: 551  
TYPE: DNA  
ORGANISM: Zea mays subsp. mexicana  
FEATURE:  
OTHER INFORMATION: Clone ID: uc-zmflteosinte90d04b1  
US-09-849-526A-27340

Query Match 76.4%; Score 16.8; DB 6; Length 551;  
Best Local Similarity 90.0%; Pred. No. 1.3e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 ttggatcacaaagttagat 22  
||| ||||| ||||| |||  
Db 264 ttggatcacaaagttagat 245

RESULT 8  
US-09-815-264-26466/C  
Sequence 26466, Application US/09815264  
GENERAL INFORMATION:  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Cao, Yongwei  
APPLICANT: Dotson, Stanton B.  
APPLICANT: Koshi, Jeffrey M.  
APPLICANT: Kovalic, David K.  
APPLICANT: Liu, Jingdong  
APPLICANT: McIninch, James  
APPLICANT: Wu, Wei  
TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof  
FILE REFERENCE: 38-21(51237)G  
CURRENT APPLICATION NUMBER: US/09/815,264  
CURRENT FILING DATE: 2001-03-23  
PRIOR APPLICATION NUMBER: US 09/620,392

PRIOR FILING DATE: 2000-07-19  
PRIOR APPLICATION NUMBER: US 09/702,134  
PRIOR FILING DATE: 2000-10-31  
NUMBER OF SEQ ID NOS: 109669  
SEQ ID NO 26466  
LENGTH: 849  
TYPE: DNA  
ORGANISM: Oryza sativa  
US-09-815-264-26466

Query Match 76.4%; Score 16.8; DB 6; Length 849;  
Best Local Similarity 90.0%; Pred. No. 1.4e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 gtggatcacaaagttagat 21  
||||| ||||| ||||| |||  
Db 256 gtggatcacaaagttagat 237

RESULT 9  
US-09-815-264-63319  
Sequence 63319, Application US/09815264  
GENERAL INFORMATION:  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Cao, Yongwei  
APPLICANT: Dotson, Stanton B.  
APPLICANT: Koshi, Jeffrey M.  
APPLICANT: Kovalic, David K.  
APPLICANT: Liu, Jingdong  
APPLICANT: McIninch, James  
APPLICANT: Wu, Wei  
TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof  
FILE REFERENCE: 38-21(51237)G  
CURRENT APPLICATION NUMBER: US/09/815,264  
CURRENT FILING DATE: 2001-03-23  
PRIOR APPLICATION NUMBER: US 09/620,392  
PRIOR FILING DATE: 2000-07-19  
PRIOR APPLICATION NUMBER: US 09/702,134  
PRIOR FILING DATE: 2000-10-31  
NUMBER OF SEQ ID NOS: 109669  
SEQ ID NO 63319  
LENGTH: 1046  
TYPE: DNA  
ORGANISM: Oryza sativa  
US-09-815-264-63319

Query Match 76.4%; Score 16.8; DB 6; Length 1046;  
Best Local Similarity 90.0%; Pred. No. 1.4e+02;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 gtggatcacaaagttagat 21  
||||| ||||| ||||| |||  
Db 791 gtggatcacaaagttagat 810

RESULT 10  
US-09-815-264-91672  
Sequence 91672, Application US/09815264  
GENERAL INFORMATION:  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Cao, Yongwei  
APPLICANT: Dotson, Stanton B.  
APPLICANT: Koshi, Jeffrey M.  
APPLICANT: Kovalic, David K.  
APPLICANT: Liu, Jingdong  
APPLICANT: McIninch, James  
APPLICANT: Wu, Wei  
TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof  
FILE REFERENCE: 38-21(51237)G  
CURRENT APPLICATION NUMBER: US/09/815,264  
CURRENT FILING DATE: 2001-03-23

US-09-815-264-76889  
Sequence 76889, Application US/09815264  
GENERAL INFORMATION:  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Cao, Yongwei  
APPLICANT: Dodson, Stanton B.  
APPLICANT: Koshi, Jeffrey M.  
APPLICANT: Kovacic, David K.  
APPLICANT: Liu, Jingdong

```

; APPLICANT: McIninch, James
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51237)G
; CURRENT APPLICATION NUMBER: US/09/815,264
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 09/620,392
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/702,134
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 109669
; SEQ ID NO: 76889
; LENGTH: 6940
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(6940)
; OTHER INFORMATION: unsure at all n locations
US-09-815-264-76889

```

```

Query Match          76.4%; Score 16.8; DB 6; Length 6940;
Best Local Similarity 90.0%; Pred. No. 1.7e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      3 ttggtcacagaatttgata 22
        ||||| || ||||| |||||
DB      3165 ttgacccaagaatttgata 3184

```

```

RESULT 15
US-09-815-264-81045/c
; Sequence 81045, Application US/09815264
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Cao, Yongwei
; APPLICANT: Dotson, Stanton B.
; APPLICANT: Koshl, Jeffrey M.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: McIninch, James
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51237)G
; CURRENT APPLICATION NUMBER: US/09/815,264
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 09/620,392
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/702,134
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 109669
; SEQ ID NO: 81045
; LENGTH: 12081
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-815-264-81045

```

```

Query Match          76.4%; Score 16.8; DB 6; Length 12081;
Best Local Similarity 90.0%; Pred. No. 1.7e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      2 gttggtcacagaatttgat 21
        ||||| ||||| ||||| |||||
DB      4391 GTTGATCTACTAGATTAGAT 4372

```

Search completed: December 15, 2001, 05:58:26  
Job time: 12269 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 03:19:57 ; Search time 2725.73 Seconds  
(without alignments)  
133.153 Million cell updates/sec

Title: US-09-380-826a-7

Perfect score: 1 tgltgatcacagatgata 22

Sequence: IDENTITY\_NUC

Scoring table: Gapop 10.0, Gapext 1.0

Searched: 1472140 segs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb-da:\*  
2: gb-hcg:\*  
3: gb-in:\*  
4: gb-om:\*  
5: gb-ov:\*  
6: gb-pat:\*  
7: gb-ph:\*  
8: gb-pl:\*  
9: gb-pr:\*  
10: gb-ro:\*  
11: gb-sts:\*  
12: gb-sv:\*  
13: gb-un:\*  
14: gb-vl:\*  
15: em-da:\*  
16: em-fun:\*  
17: em-hum:\*  
18: em-in:\*  
19: em-om:\*  
20: em-or:\*  
21: em-ov:\*  
22: em-pat:\*  
23: em-ph:\*  
24: em-pl:\*  
25: em-ro:\*  
26: em-sts:\*  
27: em-sy:\*  
28: em-un:\*  
29: em-vl:\*  
30: em-htgo\_hum:\*  
31: em-htgo\_inv:\*  
32: em-htgo\_rod:\*  
33: em-htg\_hum:\*  
34: em-htg\_inv:\*  
35: em-htg\_rod:\*  
36: em-htg\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	22	100.0	1481	1	LFU60594	U60594 Leptospira
2	20.4	92.7	354	1	AB007014	AB007014 Spirochae
3	18.8	85.5	283	1	AF003953	AF003953 Leptospir
4	18.8	85.5	288	1	LIU94974	LIU94974 Leptospira
5	18.8	85.5	288	1	LIU94975	LIU94975 Leptospira
6	18.8	85.5	288	1	LIU94976	LIU94976 Leptospira
7	18.8	85.5	288	1	LIU94977	LIU94977 Leptospira
8	18.8	85.5	288	1	LIU94978	LIU94978 Leptospira
9	18.8	85.5	288	1	LIU94979	LIU94979 Leptospira
10	18.8	85.5	353	1	AB007012	AB007012 Spirochae
11	18.8	85.5	1012	8	ATRHNS2X	M98336 Arabidopsi
12	18.8	85.5	9423	8	ATRH275979	ATRH275979 Arabidops
13	18.8	85.5	72590	8	AC003000	AC003000 Arabidops
14	18.8	85.5	121524	2	AF165146	AF165146 Homo sapi
15	18.8	85.5	123090	2	AC083961	AC083961 Homo sapi
16	18.8	85.5	182545	2	AC046176	AC046176 Homo sapi
17	18.8	85.5	200799	8	AF137379	AF137379 Nephrosel
18	18.8	85.5	200799	8	AF137379	AF137379 Nephrosel
19	18.4	83.6	1812	3	AF220067	AF220067 Drosophi
20	18.4	83.6	27365	2	AC015209	AC015209 Drosophi
21	18.4	83.6	173613	3	AC007475	AC007475 Drosophi
22	18.4	83.6	262731	3	AE003823	AE003823 Drosophi
23	17.8	80.9	1874	14	EBBPC44MCP	LA8441 Echinochloa
24	17.8	80.9	31214	8	SPCC830	AL109850 S.pombe c
25	17.8	80.9	85992	2	AB009052	AB009052 Arabidops
26	17.8	80.9	91894	2	AC026327	AC026327 Homo sapi
27	17.8	80.9	172681	9	AC026320	AC026320 Homo sapi
28	17.8	80.9	270889	2	AC055742	AC055742 Homo sapi
29	17.4	79.1	59261	8	T12M4	AC003114 Arabidops
30	17.4	79.1	81662	8	AB008265	AB008265 Arabidops
31	17.4	79.1	109476	9	AL513548	AL513548 Human DNA
32	17.4	79.1	109512	2	AC068144	AC068144 Homo sapi
33	17.4	79.1	148997	9	AC021015	AC021015 Homo sapi
34	17.4	79.1	187847	2	AC023155	AC023155 Homo sapi
35	17.2	78.2	316	1	AF352068	AF352068 Peptostre
36	17.2	78.2	462	1	AF175672	AF175672 Unculture
37	17.2	78.2	835	8	AF271231	AF271231 Albungo ca
38	17.2	78.2	1083	8	AF018567	AF018567 Unidentif
39	17.2	78.2	2213	8	LEP4CCOALB	D49367 Lithospermu
40	17.2	78.2	3608	8	SCYIL031C	D73136 S.cerevisia
41	17.2	78.2	5596	3	SPU40832	U40832 Strongyloce
42	17.2	78.2	34496	3	U41016	U41016 Caenorhabdi
43	17.2	78.2	38655	3	U41545	U41545 Caenorhabdi
44	17.2	78.2	49311	2	P26K10	AL049803 Arabidops
45	17.2	78.2	60406	9	AL590550	AL590550 Human DNA

#### ALIGNMENTS

RESULT 1  
LOCUS LFU60594 1481 bp DNA BCT 10-SEP-1998  
DEFINITION Leptospira fainei 16S ribosomal RNA gene, partial sequence.  
ACCESSION U60594  
VERSION U60594.1 GI:1408219  
KEYWORDS  
SOURCE  
ORGANISM Leptospira fainei.  
Bacteria; Spirochaetales; Leptospiraceae; Leptospira.  
REFERENCE  
1 (bases 1 to 1481)  
Perolat,P., Chappel,R.J., Adler,B., Baranton,G., Bulach,D.M.,  
Billinghurst,M.L., Letocart,M., Merlen,F. and Serrano,M.S.  
Leptospira fainei sp. nov., isolated from pigs in Australia  
Int. J. Syst. Bacteriol. 48 Pt 3, 851-858 (1998)  
TITLE JOURNAL  
MEDLINE 98404550  
REFERENCE 2 (bases 1 to 1481)  
Adler,B., Chappel,R.J., Baranton,G., Bulach,D.M.,  
Billinghurst,M.L., Letocart,M., Merlen,F., Serrano,M.S. and  
Perolat,P.

TITLE Direct Submission  
JOURNAL Submitted (12-JUN-1996) Microbiology, Monash University, Wellington  
LOCUS Rd., Clayton, VIC 3168, Australia  
FEATURES  
source  
1.1481  
/organism="Leptospira falnei"  
/strain="Hurstbridge"  
/db\_xref="taxon:48782"  
<1..>1481  
/product="16S ribosomal RNA"  
BASE COUNT 391 a 335 c 439 g 314 t 2 others  
ORIGIN

Query Match 100.0%; Score 22; DB 1; Length 1481;  
Best Local Similarity 100.0%; Pred. No. 2;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttgatcacaaagtgtgata 22  
|||||  
Db 155 TGTGATCACAGATTGTGATA 176

RESULT 2  
LOCUS AB007014 354 bp DNA BCT 13-OCT-1997  
DEFINITION Spirochaeta sp. 16S rRNA gene, partial sequence.  
ACCESSION AB007014  
VERSION AB007014.1 GI:2516255  
KEYWORDS 16S ribosomal RNA.  
SOURCE Spirochaeta sp. (sub\_species:Freshwater obligate oligotroph,  
strain:FO-95) DNA.  
ORGANISM Spirochaeta sp.  
REFERENCE 1 (bases 1 to 354)  
AUTHORS Bacteria; Spirochaetales; Spirochaetaceae; Spirochaeta.  
TITLE Shln.M.-S.  
JOURNAL Direct Submission  
SUBMITTED (05-SEP-1997) to the DDBJ/EMBL/Genbank databases. M.-Sun  
Shln, Laboratory of Marine Molecular Microbiology, Faculty of  
Agriculture, Kyoto University, Oiwakecho, Kitashirakawa, Sakyo-ku,  
Kyoto, Kyoto 606-01, Japan (E-mail:misun@kais.kyoto-u.ac.jp,  
Tel:075-753-6224, Fax:075-753-6226)  
2 (sites)  
REFERENCE Shln.M., Yoshinaga,I., Uchida,A. and Ishida,Y.  
AUTHORS Phylogenetic analysis by 16S rRNA gene sequencing of obligate  
TITLE oligotrophs isolated from the northern basin of Lake Biwa  
(mesotrophic Lake)  
JOURNAL Unpublished (1997)  
FEATURES  
source  
Location/Qualifiers  
1..354  
/organism="Spirochaeta sp."  
/strain="FO-95"  
/sub\_species="Freshwater obligate oligotroph"  
/db\_xref="taxon:28185"  
<1..>354  
rRNA /product="16S rRNA"  
BASE COUNT 91 a 79 c 111 g 73 t  
ORIGIN

Query Match 92.7%; Score 20.4; DB 1; Length 354;  
Best Local Similarity 95.5%; Pred. No. 13;  
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 tgttgatcacaaagtgtgata 22  
|||||  
Db 133 TGTGATCACAGATTGTGATA 154

RESULT 3  
LOCUS AF003953 283 bp DNA BCT 30-MAY-1998  
DEFINITION Leptospira interrogans strain 48/95 16S ribosomal RNA gene, partial

sequence.  
ACCESSION AF003953  
VERSION AF003953.1 GI:3169306  
KEYWORDS  
SOURCE  
ORGANISM Leptospira interrogans.  
REFERENCE 1 (bases 1 to 283)  
AUTHORS Bacteria; Spirochaetales; Leptospiraceae; Leptospira.  
TITLE Moo,T.H.S., Patel,B.K.C., Cinco,M., Smythe,L.D., Symonds,M.,  
Norris,M. and Dohnt,M.  
JOURNAL Identification of Leptonema by real-time homogeneous assay of rapid  
cycle PCR product  
REFERENCE 2 (bases 1 to 283)  
AUTHORS Unpublished  
JOURNAL  
TITLE Moo,T.H.S., Patel,B.K.C., Cinco,M., Smythe,L.D., Symonds,M.,  
Norris,M. and Dohnt,M.  
JOURNAL Direct Submission  
SUBMITTED (14-MAY-1997) School of Science, Griffith University,  
Brisbane, QLD 4111, Australia  
FEATURES  
source  
Location/Qualifiers  
1..283  
/organism="Leptospira interrogans"  
/strain="48/95"  
/db\_xref="taxon:173"  
<1..>283  
rRNA /product="16S rRNA"  
BASE COUNT 74 a 60 c 88 g 61 t  
ORIGIN

Query Match 85.5%; Score 18.8; DB 1; Length 283;  
Best Local Similarity 90.9%; Pred. No. 78;  
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tgttgatcacaaagtgtgata 22  
|||||  
Db 136 TATTGATCACAGCATTTGATA 157

RESULT 4  
LOCUS L1U94974 288 bp DNA BCT 01-JAN-1998  
DEFINITION Leptospira inada1 16S ribosomal RNA gene, partial sequence.  
ACCESSION U94974  
VERSION U94974.1 GI:2735446  
KEYWORDS  
SOURCE  
ORGANISM Leptospira inada1.  
REFERENCE 1 (bases 1 to 288)  
AUTHORS Bacteria; Spirochaetales; Leptospiraceae; Leptospira.  
TITLE Woo,T.H.S., Smythe,L.D., Symonds,M., Norris,M., Dohnt,M.,  
Brenner,D.J. and Patel,B.K.C.  
JOURNAL Identification of Leptospira inada1 by continuously monitoring  
REFERENCE 2 (bases 1 to 288)  
AUTHORS Unpublished  
JOURNAL  
TITLE Woo,T.H.S., Smythe,L.D., Symonds,M., Norris,M., Dohnt,M.,  
Brenner,D.J. and Patel,B.K.C.  
JOURNAL Direct Submission  
SUBMITTED (24-MAR-1997) School of Science, Griffith University,  
Brisbane, QLD 4111, Australia  
FEATURES  
source  
Location/Qualifiers  
1..288  
/organism="Leptospira inada1"  
/strain="79/95"  
/db\_xref="taxon:29506"  
<1..>288  
rRNA /product="16S ribosomal RNA"  
BASE COUNT 77 a 61 c 88 g 62 t  
ORIGIN

Query Match 85.5%; Score 18.8; DB 1; Length 288;

Best Local Similarity 90.9%; Pred. No. 78;  
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tgttgatcacaaagatttgata 22  
1 |||||  
Db 141 TATTGGATCAGAGATTTGATA 162

## RESULT 5

LIU94975 288 bp DNA BCT 01-JAN-1998  
LOCUS Leptospira inadai 16S ribosomal RNA gene, partial sequence.  
DEFINITION U94975  
ACCESSION U94975  
VERSION U94975.1 GI:2735447

SOURCE  
ORGANISM Leptospira inadai.  
Bacteria; Spirochaetales; Leptospiraceae; Leptospira.

REFERENCE  
AUTHORS Woo, T.H.S., Smythe, L.D., Symonds, M., Norris, M., Dohnl, M.,  
Brenner, D.J. and Patel, B.K.C.

TITLE Identification of Leptospira inadai by continuously monitoring  
fluorescence during rapid cycle PCR  
Unpublished

JOURNAL  
REFERENCE 2 (bases 1 to 288)  
AUTHORS Woo, T.H.S., Smythe, L.D., Symonds, M., Norris, M., Dohnl, M.,  
Brenner, D.J. and Patel, B.K.C.

TITLE Direct Submission  
JOURNAL Submitted (24-MAR-1997) School of Science, Griffith University,  
Brisbane, QLD 4111, Australia

FEATURES  
source 1..288  
/organism="Leptospira inadai"  
/strain="68/94"  
/db\_xref="taxon:29506"

rRNA  
/product="16S ribosomal RNA"  
ORIGIN  
77 a 61 c 88 g 62 t

BASE COUNT  
ORIGIN

Query Match 85.5%; Score 18.8; DB 1; Length 288;  
Best Local Similarity 90.9%; Pred. No. 78;  
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tgttgatcacaaagatttgata 22  
1 |||||  
Db 141 TATTGGATCAGAGATTTGATA 162

## RESULT 6

LIU94976 288 bp DNA BCT 01-JAN-1998  
LOCUS Leptospira inadai 16S ribosomal RNA gene, partial sequence.  
DEFINITION U94976  
ACCESSION U94976  
VERSION U94976.1 GI:2735448

SOURCE  
ORGANISM Leptospira inadai.  
Bacteria; Spirochaetales; Leptospiraceae; Leptospira.

REFERENCE  
AUTHORS Woo, T.H.S., Smythe, L.D., Symonds, M., Norris, M., Dohnl, M.,  
Brenner, D.J. and Patel, B.K.C.

TITLE Identification of Leptospira inadai by continuously monitoring  
fluorescence during rapid cycle PCR  
Unpublished

JOURNAL  
REFERENCE 2 (bases 1 to 288)  
AUTHORS Woo, T.H.S., Smythe, L.D., Symonds, M., Norris, M., Dohnl, M.,  
Brenner, D.J. and Patel, B.K.C.

TITLE Direct Submission  
JOURNAL Submitted (24-MAR-1997) School of Science, Griffith University,  
Brisbane, QLD 4111, Australia

FEATURES  
Location/Qualifiers

source 1..288  
/organism="Leptospira inadai"  
/strain="346/95"  
/db\_xref="taxon:29506"

rRNA  
/product="16S ribosomal RNA"  
ORIGIN  
77 a 61 c 88 g 62 t

BASE COUNT  
ORIGIN

Query Match 85.5%; Score 18.8; DB 1; Length 288;  
Best Local Similarity 90.9%; Pred. No. 78;  
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tgttgatcacaaagatttgata 22  
1 |||||  
Db 141 TATTGGATCAGAGATTTGATA 162

## RESULT 7

LIU94977 288 bp DNA BCT 01-JAN-1998  
LOCUS Leptospira inadai 16S ribosomal RNA gene, partial sequence.  
DEFINITION U94977  
ACCESSION U94977  
VERSION U94977.1 GI:2735449

SOURCE  
ORGANISM Leptospira inadai.  
Bacteria; Spirochaetales; Leptospiraceae; Leptospira.

REFERENCE  
AUTHORS Woo, T.H.S., Smythe, L.D., Symonds, M., Norris, M., Dohnl, M.,  
Brenner, D.J. and Patel, B.K.C.

TITLE Identification of Leptospira inadai by continuously monitoring  
fluorescence during rapid cycle PCR  
Unpublished

JOURNAL  
REFERENCE 2 (bases 1 to 288)  
AUTHORS Woo, T.H.S., Smythe, L.D., Symonds, M., Norris, M., Dohnl, M.,  
Brenner, D.J. and Patel, B.K.C.

TITLE Direct Submission  
JOURNAL Submitted (24-MAR-1997) School of Science, Griffith University,  
Brisbane, QLD 4111, Australia

FEATURES  
source 1..288  
/organism="Leptospira inadai"  
/strain="268/95"  
/db\_xref="taxon:29506"

rRNA  
/product="16S ribosomal RNA"  
ORIGIN  
77 a 61 c 88 g 62 t

Query Match 85.5%; Score 18.8; DB 1; Length 288;  
Best Local Similarity 90.9%; Pred. No. 78;  
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tgttgatcacaaagatttgata 22  
1 |||||  
Db 141 TATTGGATCAGAGATTTGATA 162

## RESULT 8

LIU94978 288 bp DNA BCT 01-JAN-1998  
LOCUS Leptospira inadai 16S ribosomal RNA gene, partial sequence.  
DEFINITION U94978  
ACCESSION U94978  
VERSION U94978.1 GI:2735450

SOURCE  
ORGANISM Leptospira inadai.  
Bacteria; Spirochaetales; Leptospiraceae; Leptospira.

REFERENCE  
AUTHORS Woo, T.H.S., Smythe, L.D., Symonds, M., Norris, M., Dohnl, M.,  
Brenner, D.J. and Patel, B.K.C.

TITLE	Brenner,D.J., and Patel,B.K.C.
JOURNAL REFERENCE	Identification of Leptospira inadai by continuously monitoring fluorescence during rapid cycle PCR
AUTHORS	Unpublished
TITLE	2 (bases 1 to 288)
JOURNAL	Woo,T.H.S., Smythe,L.D., Symonds,M., Norris,M., Dohnt,M., Brenner,D.J. and Patel,B.K.C.
FEATURES	Direct Submission Submitted (24-MAR-1997) School of Science, Griffith University, Brisbane, QLD 4111, Australia
SOURCE	Location/Qualifiers 1..288
rRNA	/organism="Leptospira inadai" /strain="218/95" /db_xref="taxon:29506" <1..>288 /product="16S ribosomal RNA"
BASE COUNT	77 a 61 c 88 g 62 t
ORIGIN	
Query Match	85.5%; Score 18.8; DB 1; Length 288;
Best Local Similarity	90.9%; Pred. No. 78;
Matches	20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY	1 tcttgatcacagaattgata 22       Db 141 TATTGGATCACAGGATTGTGATA 162
RESULT 9	
LOCUS	L1U94979 288 bp DNA BCT 01-JAN-1998
DEFINITION	Leptospira inadai 16S ribosomal RNA gene, partial sequence.
ACCESSION	U94979
VERSION	U94979.1 GI:2735451
KEYWORDS	
ORGANISM	Leptospira inadai.
REFERENCE	Leptospira inadai. Bacteria; Spirochaetales; Leptospiaceae; Leptospira.
AUTHORS	1 (bases 1 to 288) Woo,T.H.S., Smythe,L.D., Symonds,M., Norris,M., Dohnt,M., Brenner,D.J. and Patel,B.K.C.
TITLE	Identification of Leptospira inadai by continuously monitoring fluorescence during rapid cycle PCR
JOURNAL REFERENCE	Unpublished
AUTHORS	2 (bases 1 to 288) Woo,T.H.S., Smythe,L.D., Symonds,M., Norris,M., Dohnt,M., Brenner,D.J. and Patel,B.K.C.
TITLE	Direct Submission
JOURNAL	Submitted (24-MAR-1997) School of Science, Griffith University, Brisbane, QLD 4111, Australia
FEATURES	Location/Qualifiers 1..288
SOURCE	/organism="Leptospira inadai" /strain="1078 VRI" /db_xref="taxon:29506" <1..>288 /product="16S ribosomal RNA"
rRNA	
BASE COUNT	77 a 61 c 88 g 62 t
ORIGIN	
Query Match	85.5%; Score 18.8; DB 1; Length 288;
Best Local Similarity	90.9%; Pred. No. 78;
Matches	20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY	1 tcttgatcacagaattgata 22       Db 141 TATTGGATCACAGGATTGTGATA 162
RESULT 10	

AB007012	AB007012	353 bp	DNA	BCP	13-OCT-1997
LOCUS	Spirochaeta sp. 16S rRNA gene, partial sequence.				
DEFINITION	Spirochaeta sp. 16S rRNA gene, partial sequence.				
ACCESSION	AB007012				
VERSION	AB007012.1 GI:2516253				
KEYWORDS	16S ribosomal RNA.				
SOURCE	Spirochaeta sp. (sub_species:Freshwater obligate oligotroph, strain:SO-104) DNA.				
ORGANISM	Spirochaeta sp.				
REFERENCE	Bacteria: Spirochaetales; Spirochaetaceae; Spirochaeta.				
AUTHORS	1 (bases 1 to 353)				
TITLE	Shin,M.,S.				
JOURNAL	Direct Submission				
FEATURES	Submitted (05-SEP-1997) to the DDBJ/EMBL/Genbank databases. Mi-Sun Shin, Laboratory of Marine Molecular Microbiology, Faculty of Agriculture, Kyoto University; Oiwakecho, Kitashirakawa, Sakyo-ku, Kyoto, Kyoto 306-01, Japan (E-mail:mtsun@kais.kyoto-u.ac.jp, Tel:075-753-6224, Fax:075-753-6226)				
REFERENCE	2 (sites)				
AUTHORS	Shin,M., Yoshinaga,I., Uchida,A. and Ishida,Y.				
TITLE	Phylogenetic analysis by 16S rRNA gene sequencing of Obligate Oligotrophs Isolated from the northern basin of Lake Biwa (Mesotrophic Lake)				
JOURNAL	unpublished (1997)				
FEATURES	Location/Qualifiers				
source	1..353				
	/organism="Spirochaeta sp."				
	/strain="SO-104"				
	/sub_species="Freshwater obligate oligotroph"				
	/db_xref="taxon:28185"				
	<1..>353				
rrna	/product="16S rRNA"				
BASE COUNT	90 a 80 c 108 g 75 t				
ORIGIN					
Query Match	85.5%; Score 18.8; DB 1; Length 353;				
Best Local Similarity	90.9%; Pred. NO. 77;				
Matches	20: Conservative 0; Mismatches 2; Indels 0; Gaps 0.				
OY	1 ttttgatcacaaagatttgata 22				
Db	132 TGTGGATCACCAGATCTGATA 153				
RESULT 11					
ATHRNS2X	ATHRNS2X 1012 bp mRNA PLN 30-OCT-1994				
LOCUS	Arabidopsis thaliana ribonuclease (RNS2) mRNA, complete cds.				
DEFINITION	Arabidopsis thaliana ribonuclease (RNS2) mRNA, complete cds.				
ACCESSION	M98336				
VERSION	M98336.1 GI:289209				
KEYWORDS	ribonuclease.				
SOURCE	Arabidopsis thaliana cDNA to mRNA.				
ORGANISM	Arabidopsis thaliana				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.				
REFERENCE	1 (bases 1 to 1012)				
AUTHORS	Taylor,C.B., Bariola,P.A., delCardayre,S.B., Raines,R.T. and Green,P.J.				
TITLE	RNS2: a senseless-associated RNase of Arabidopsis that diverged from the S-Rnases before speciation				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 90 (11), 5118-5122 (1993)				
MEDLINE	93281708				
FEATURES	Location/Qualifiers				
source	1..1012				
	/organism="Arabidopsis thaliana"				
	/db_xref="taxon:3702"				
gene	16..795				
	/gene="RNS2"				
cds	16..795				
	/gene="RNS2"				
	/codon_start=1				



Genes were identified by a combination of three methods: Gene prediction programs including GRAIL (<http://arthur.epm.ornl.gov/pub/xbgrail>), GeneFinder (Phil Green, University of Washington), GenScan (Chris Burge, <http://genomic.stanford.edu/GENSCAN.html>), and NetPlantGene (<http://www.cds.dtu.dk/services/NetGene2/>), searches of the complete sequence against a peptide database and plant EST databases at TIGR, and manual curations based on those analyses. Annotated genes are named to indicate the level of evidence for their annotation. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as 'unknown' proteins. Genes without protein or EST similarity, that are predicted by two or more gene prediction programs over most of their length are annotated as 'hypothetical' proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, <http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats were identified by RepeatMasker (Arian Smit, <http://ftp.genome.washington.edu/RM/RepeatMasker.html>). Genes are numbered from the top to bottom of the chromosome.

We thank the CSHL/MashU/ABI consortium for sequencing BAC clones FBP23, F5J6, T17A5, and T13L16, the ESSA group for sequencing clone F13D4, and Scott Jackson, Jiming Jiang, Klaus Meyer, Eric Richards and Satoshi Tabata for helpful assistance. In addition, we would like to thank the TIGR Bioinformatics Department, especially Lixin Zhou, Hanli Khatak, Michael E. Heaney, Lily Fu, Feng Liang, Jeremy Peterson, Michael Holmes, and Delwood Richardson for software and database support.

This work was supported by the National Science Foundation, Department of Energy and the US Department of Agriculture.

Address all correspondence to: [ateligr.org](mailto:ateligr.org).

FEATURES	Address all correspondence to: <a href="mailto:ateligr.org">ateligr.org</a> .
source	Location/Qualifiers
	1. 72590
	/organism="Arabidopsis thaliana"
	/cultivar="Columbia"
	/db_xref="taxon:3702"
	/chromosome="11"
misc_feature	1. .>72590
	/note="Sequence from clone T517"
repeat_region	complement(130. .198)
	/rpt_family="(CAAA)n"
repeat_region	complement(1000. .1024)
	/rpt_family="POLY_A"
repeat_region	complement(11266. .1291)
	/rpt_family="(CAAA)n"
repeat_region	complement(2759. .2862)
	/rpt_family="POLY_A"
gene	complement(<4584. .>5909)
	/gene="At2g39710"
	complement(<4584. .>5909)
	/gene="At2g39710"
	/note="T517.1: similar to GP122450121gn1 PID e3270001 297341"
CDS	complement(4584. .5909)
	/gene="At2g39710"
	/note="unknown protein"
	/codon_start=1
	/protein_id="AAB87120.1"
	/db_xref="GI:2642153"
	/translation="MDSSTSSLSLSKFLRISVLLIFLPTCKSTSTNQLLSLK TOLKPOSSDKLSFRHNVTLLVTLAVGPPONISMVLDTGSELSMLCKSPNLSYF NPVSSSTYSPGSCPTCRTRDLPIPASCDPTLCHVAISYADATISIGNLHET FVTSVTRPGTLFGCMDSGLSSNSEPAKSGMLNGRSLSYNOIGFESEKSTIGS SDSSGFLLDASTSWLGPIDYTPPLVLDSTPLPFDRAVTVQLEGTRVSKILSLK SVFVDDHTGAGTQVDSGOTFLMGVYTLAKNEFTQTKSVLRVLDDEDPVCFQGT DLCKVSGTTPNFSGLPVMSLMRGAMSVSGKLLYRVNGASGEQEEVYCTFGN SDLGIEAFVIGHHQQNVMMFEDLAKSRVGFAGNVRCDLASORLGLND"
MRNA	complement(<10039. .11304)
	/gene="At2g39720"
	complement(<10039. .11304)
	/gene="At2g39720"
	complement(13039. .11244)
	/gene="At2g39720"
	/note="unknown protein"
	/codon_start=1
	/protein_id="AAB87121.1"
	/db_xref="GI:2642154"
	/translation="MASGSYKCYSCSRFVWSDISCPDCCGGFLELTOEPLDFPSPD SFPTTPTTQHRSPTRPPSSSSSTPSPASMAHDSNPTPIYTRTNSKNSPNVIYLRG SAAASSDVSGDLDRSAFQMYDDGDTSGLPSPMTEFLGSGFURLDQISOIE LMTNRNLRCSEHPSPAKSAIEALPLEIDPTHLSDSQSHCAVCKENPLKSSAPEM CHNHPDCLPMLAIRNCPYCRHLEPDEDLTDGTALITVTAETAEEDSAGLT IMRIGCGGFAVRIGGMRGDRMPVYVTEVDGRLGERLPRVAVMSRRGGRCGC GSRERCGFAGATMRLFCGFSSSSGSIATAAASSGSGRIYVTRTRTSFSMFVASS SSRRRWLA"
repeat_region	complement(11453. .11512)
	/rpt_family="POLY_A"
MRNA	complement(join(13323. .13599,13724. .14005,14087. .14559, 14651. .14736,14836. .14941,15033. .15349,15834. .16055))
	/gene="At2g39730"
mRNA	complement(join(13323. .13599,13713. .14005,14087. .14559, 14651. .14736,14836. .14941,15033. .15349,15834. .16055))
	/gene="At2g39730"
gene	complement(13323. .16055)
	/gene="At2g39730"
	/note="T517.3"
	13364. .13461
repeat_region	/rpt_family="POLY_A"
CDS	complement(join(13484. .13599,13724. .14005,14087. .14559, 14651. .14736,14836. .14941,15033. .15349,15834. .15878))
	/gene="At2g39730"
	/note="alternative splice form 1 GP1166834 M86720"
	/codon_start=1
	/product="Rubisco activase"
	/protein_id="AAB87122.1"
	/db_xref="GI:2642155"
	/translation="MAAAVSTGAIRAPLPLSLGSGSAGVASAPSTFLCKVYVSFR AOSNKSNGSFYKLAVERDQOTGDGRRGRLAYTSDQODITRGKMGVSVQAPAGT GTHAVLSYSEYVSGLRQYNDNMMDGFTIAPFMKLVVHTIKNFLLPNTKVPIL LGIMGCGGKSGFQCELVMAKMGINPTMSAGLESGNAGEPAKLIRORYREAADLIK KGKMKCLFINDLDAGAGRMGKTQYTNMOMVATLMNIADNPTNQLPGMYKKEENA RVPIITCGNDFSTLYAPLIRDRGMEKRYVAPPRPREDTIGYCKGTFRTDKIKDEDIVLY DQPPGOSIDFEGALRAVRVDDEVKRFVESIGVKRIGRLNSRSGPVPFQPPMTEYK LMEYGNMLVMEQENVKRVQALAEYTLQAAIGDANADAIGKFTYKTEKEPEPK"
	complement(11761. .17639)
repeat_region	/rpt_family="(TA)n"
MRNA	complement(<18783. .18921,19188. .19352,19496. .19674,19766. .19861, 19965. .20034,20131. .20261,20593. .>21237)
	/gene="At2g39740"
gene	<18783. .>21237
	/gene="At2g39740"
	/note="T517.4: predicted by genscan"
CDS	complement(join(18783. .18921,19188. .19352,19496. .19674,19766. .19861, 19965. .20034,20131. .20261,20593. .21237))

/gene="At2g39740"  
/note="hypothetical protein"  
/codon\_start=1  
/protein\_id="AAB87123.1"  
/db\_xref="gi:2642156"  
/translation="MSRNPFLDPTLQELIIVKIPPRARDRTITVVDLQRLVQSVLC  
LRGATVDFGFSVNLFTRMGDLISVDFSGSSILFGKSKOKOTLGLRLARASG  
LWYLOPFGHARVILKAVSGHORISDIDNLGILKSRFLMISLIDGRFDLV  
LTKEMAKAHNTNDKGTGFSYSLIVYHPCVCVPLPPLVYIKRSVVDLTG  
RKTRBESIAQVTANINAFKSEKRSKSVRSLSLVSFAKVEDPEQPYNAARSY  
RRLNDRIAQVEQITSRRLVSECNRNSIIGILTGQHIDSLYRTISLSPQHMANHNY  
RLHGQARPOQMOQMOQMSOSYNTPNPMPPLTOSRPOQMTONTNPNLQGPVVG  
QTPVITOTQOQSPYKSGNRPLKNTSAGSSQNGHIGKPSGHMNGVNSARPAYTG  
VNSARPSKIPRSOGQIWRPRRQ"  
complement(join(21380..21726,21839..21916,22124..22246,  
22328..22585,22834..22993,23212..23322,23481..23751,  
23847..23996,24089..>24814))  
/gene="At2g39750"  
complement(21380..>24814)

gene  
Query Match 85.5%; Score 18.8; DB 8; Length 72590;  
Best Local Similarity 90.9%; Pred. No. 72;  
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 tcttgatcacagaatttgata 22  
||||| |||||||||  
Db 35807 TCTTGATCACAGAATTTCACA 35828

RESULT 14  
AF165146 121524 bp DNA HTG 07-JUN-2001  
LOCUS Homo sapiens chromosome 8 clone CTA-397H3 map 8q12-8q13, \*\*\*  
DEFINITION SEQUENCING IN PROGRESS \*\*\*, 7 unordered pieces.  
ACCESSION AF165146  
VERSION AF165146.3 GI:14327840  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_FULLTOP.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
AUTHORS 1 (bases 1 to 121524)  
Schillhabel,M.B., Baumgart,C., Blechschmidt,K., Detle,M., Jahn,N.,  
Lehmann,R., Menzel,U., Polley,A., Reischald,K., Schudy,A.,  
Siddiqui,R., Taudien,S., Wen,G., Schlegelberger,B., Siebert,R.,  
Rosenthal,A. and Platzer,M.  
TITLE Chromosome 8 genomic sequence  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 121524)  
Schudy,A., Blechschmidt,K., Schillhabel,M., Baumgart,C., Menzel,U.,  
Weber,J., Schattevoy,R. and Rosenthal,A.  
TITLE Direct Submission  
JOURNAL Submitted (06-JUL-1999) Genome Analysis, Institute of Molecular  
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany  
COMMENT On Jun 7, 2001 this sequence version replaced gi:8151945.  
----- Genome Center  
Center: Institute of Molecular Biotechnology  
Center code: IMB  
Web site: http://genome.imb-jena.de/  
Contact: gscj-submit@genome.imb-jena.de  
----- Project Information  
Center project name: H171  
Center clone name: CTA-397H3  
----- Summary Statistics  
Sequencing Vector: M13; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 115981 bases at least Q40  
Consensus quality: 117981 bases at least Q30  
Consensus quality: 119332 bases at least Q20  
Quality coverage: 9.57 in Q20 bases; sum-of-contigs

\* NOTE: This is a 'working draft' sequence. It currently

\* consists of 7 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
1 1187: contig of 1187 bp in length  
\* 1188 1287: gap of unknown length  
\* 1288 3530: contig of 2243 bp in length  
\* 3531 3630: gap of unknown length  
\* 3631 21624: contig of 17994 bp in length  
\* 21625 21724: gap of unknown length  
\* 21725 43656: contig of 21932 bp in length  
\* 43657 43755: gap of unknown length  
\* 43757 61383: contig of 17627 bp in length  
\* 61384 61483: gap of unknown length  
\* 61484 85318: contig of 23835 bp in length  
\* 85319 85419: gap of unknown length  
\* 85419 121524: contig of 36106 bp in length.

FEATURES  
source  
1..121524  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="8"  
/map="8q12-8q13"  
/clone="CTA-397H3"  
BASE COUNT 34959 a 25555 c 25579 g 34831 t 600 others  
ORIGIN

Query Match 85.5%; Score 18.8; DB 2; Length 121524;  
Best Local Similarity 90.9%; Pred. No. 71;  
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Oy 1 tcttgatcacagaatttgata 22  
||||| |||||||||  
Db 7103 TTTTGATCACAGAATTTCGTA 7124

RESULT 15  
AC083961 129090 bp DNA HTG 15-FEB-2001  
LOCUS Homo sapiens chromosome 8 clone RP11-22E14 map 8, \*\*\* SEQUENCING IN  
DEFINITION PROGRESS \*\*\*, 67 unordered pieces.  
ACCESSION AC083961  
VERSION AC083961.2 GI:12831382  
KEYWORDS HTG; HTGS\_PHASE1.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
AUTHORS 1 (bases 1 to 129090)  
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,  
Anderson,S., Barra,N., Bastien,V., Beda,F., Boguslavsky,L.,  
Bouckgalter,B., Brown,A., Burkett,G., Campoliano,A., Castle,A.,  
Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,P.,  
Deatellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,  
Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M.,  
Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L.,  
Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., LaRoque,K.,  
Lamazares,R., Landers,T., Lehotzky,J., Levine,R., Liu,C., Liu,G.,  
Macdonald,P., Marquis,N., McCarthy,M., McEvan,P., McKernan,K.,  
McPheters,R., Meldrum,J., Meneus,L., Mihova,T., Mlenga,V.,  
Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T.,  
O'Donnell,P., O'Neill,D., Olivarez,T.M., Oliver,J., Peterson,K.,  
Pierre,N., Pisanl,C., Pollara,V., Raymond,C., Rieback,M., Riley,R.,  
Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P.,  
Sougnuez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,

TITLE  
JOURNAL  
COMMENT

Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,  
Tirrell, A., Travers, M., Trigglio, D., Vassiliev, H., Viel, R., Vo, A.,  
Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J.,  
Zimmer, A. and Zody, M.

## Direct Submission

Submitted (08-OCR-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Feb 15, 2001 this sequence version replaced g1:10717228.  
All repeats were identified using RepeatMasker:

Smil, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

## Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [submissions@genome.wi.mit.edu](mailto:submissions@genome.wi.mit.edu)

## Project Information

Center project name: L11257

Center clone name: 22\_E\_14

NOTE: This is a 'working draft' sequence. It currently  
consists of 67 contigs. The true order of the pieces  
is not known and their order in this sequence record is  
arbitrary. Gaps between the contigs are represented as  
runs of N, but the exact sizes of the gaps are unknown.  
This record will be updated with the finished sequence  
as soon as it is available and the accession number will  
be preserved.

1 1369: contig of 1369 bp in length  
\* 1370 1469: gap of 100 bp  
\* 1470 2279: contig of 810 bp in length  
\* 2280 2379: gap of 100 bp  
\* 2380 3192: contig of 813 bp in length  
\* 3193 3292: gap of 100 bp  
\* 3293 3539: contig of 247 bp in length  
\* 3540 3639: gap of 100 bp  
\* 3640 4530: contig of 891 bp in length  
\* 4531 4630: gap of 100 bp  
\* 4631 5236: contig of 606 bp in length  
\* 5237 5336: gap of 100 bp  
\* 5337 6055: contig of 719 bp in length  
\* 6056 6155: gap of 100 bp  
\* 6156 6653: contig of 498 bp in length  
\* 6654 6753: gap of 100 bp  
\* 6754 7388: contig of 635 bp in length  
\* 7389 7488: gap of 100 bp  
\* 7489 8145: contig of 657 bp in length  
\* 8146 8245: gap of 100 bp  
\* 8246 9151: contig of 906 bp in length  
\* 9152 9251: gap of 100 bp  
\* 9252 9962: contig of 711 bp in length  
\* 9963 10062: gap of 100 bp  
\* 10063 10776: contig of 714 bp in length  
\* 10777 10876: gap of 100 bp  
\* 10877 11474: contig of 598 bp in length  
\* 11475 11574: gap of 100 bp  
\* 11575 12639: contig of 1065 bp in length  
\* 12640 12739: gap of 100 bp  
\* 12740 13952: contig of 1213 bp in length  
\* 13953 14052: gap of 100 bp  
\* 14053 14838: contig of 786 bp in length  
\* 14839 14938: gap of 100 bp  
\* 14939 15849: contig of 911 bp in length  
\* 15850 15949: gap of 100 bp  
\* 15950 18036: contig of 2087 bp in length  
\* 18037 18136: gap of 100 bp  
\* 18137 19282: contig of 1146 bp in length  
\* 19283 19382: gap of 100 bp  
\* 19383 20437: contig of 1055 bp in length  
\* 20438 20537: gap of 100 bp  
\* 20538 21159: contig of 622 bp in length  
\* 21160 21259: gap of 100 bp  
\* 21260 22494: contig of 1235 bp in length

22495 22594: gap of 100 bp  
\* 22595 23478: contig of 884 bp in length  
\* 23479 23578: gap of 100 bp  
\* 23579 24600: contig of 1022 bp in length  
\* 24601 24700: gap of 100 bp  
\* 24701 24891: contig of 191 bp in length  
\* 24892 24991: gap of 100 bp  
\* 24992 26183: contig of 1192 bp in length  
\* 26184 26283: gap of 100 bp  
\* 26284 27662: contig of 1579 bp in length  
\* 27663 27962: gap of 100 bp  
\* 27963 28581: contig of 619 bp in length  
\* 28582 28681: gap of 100 bp  
\* 28682 29511: contig of 830 bp in length  
\* 29512 29611: gap of 100 bp  
\* 29612 31038: contig of 1427 bp in length  
\* 31039 31138: gap of 100 bp  
\* 31139 32582: contig of 1444 bp in length  
\* 32583 32682: gap of 100 bp  
\* 32683 33540: contig of 858 bp in length  
\* 33541 33640: gap of 100 bp  
\* 33641 35263: contig of 1623 bp in length  
\* 35264 35363: gap of 100 bp  
\* 35364 37036: contig of 1673 bp in length  
\* 37037 37136: gap of 100 bp  
\* 37137 38963: contig of 1827 bp in length  
\* 38964 39063: gap of 100 bp  
\* 39064 40807: contig of 1744 bp in length  
\* 40808 40907: gap of 100 bp  
\* 40908 42896: contig of 1989 bp in length  
\* 42897 42996: gap of 100 bp  
\* 42997 44598: contig of 1602 bp in length  
\* 44599 44698: gap of 100 bp  
\* 44699 46343: contig of 1645 bp in length  
\* 46344 46443: gap of 100 bp  
\* 46444 47574: contig of 1131 bp in length  
\* 47575 47674: gap of 100 bp  
\* 47675 50217: contig of 2543 bp in length  
\* 50218 50317: gap of 100 bp  
\* 50318 51338: contig of 1021 bp in length  
\* 51339 51438: gap of 100 bp  
\* 51439 53743: contig of 2305 bp in length  
\* 53744 53843: gap of 100 bp  
\* 53844 55260: contig of 1417 bp in length  
\* 55261 55360: gap of 100 bp  
\* 55361 57843: contig of 2483 bp in length  
\* 57844 57943: gap of 100 bp  
\* 57944 59865: contig of 1922 bp in length  
\* 59866 59965: gap of 100 bp  
\* 59966 62260: contig of 2295 bp in length  
\* 62261 62360: gap of 100 bp  
\* 62361 64120: contig of 1760 bp in length  
\* 64121 64220: gap of 100 bp  
\* 64221 65979: contig of 1759 bp in length  
\* 65980 66079: gap of 100 bp  
\* 66080 68712: contig of 2633 bp in length  
\* 68713 68812: gap of 100 bp  
\* 68813 71145: contig of 2333 bp in length  
\* 71146 71245: gap of 100 bp  
\* 71246 74155: contig of 2910 bp in length  
\* 74156 74255: gap of 100 bp  
\* 74256 77041: contig of 2786 bp in length  
\* 77042 77141: gap of 100 bp  
\* 77142 80685: contig of 3544 bp in length  
\* 80686 80785: gap of 100 bp  
\* 80786 83825: contig of 3040 bp in length  
\* 83826 83925: gap of 100 bp  
\* 83926 87721: contig of 3796 bp in length  
\* 87722 87821: gap of 100 bp  
\* 87822 92728: contig of 4907 bp in length  
\* 92729 92828: gap of 100 bp  
\* 92829 96652: contig of 3824 bp in length  
\* 96653 96752: gap of 100 bp



```

* 96753 99906: contig of 3154 bp in length
* 99907 100006: gap of 100 bp
* 100007 102910: contig of 2904 bp in length
* 102911 103010: gap of 100 bp
* 103011 107188: contig of 4178 bp in length
* 107189 107288: gap of 100 bp
* 107289 111047: contig of 3759 bp in length
* 111048 111147: gap of 100 bp
* 111148 116741: contig of 5594 bp in length
* 116742 116841: gap of 100 bp
* 116842 121990: contig of 5149 bp in length
* 121991 122090: gap of 100 bp
* 122091 126622: contig of 4532 bp in length
* 126623 126722: gap of 100 bp
* 126723 129090: contig of 2368 bp in length.

```

FEATURES  
Source

```

1. 129090
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8"
/clone="RP11-22E14"
/clone_lib="RPCT-11 Human Male BAC"

```

```

Query Match      85.5%; Score 18.8; DB 2; Length 129090;
Best Local Similarity 90.9%; Pred. No. 71;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 tgttgatcacagaattgata 22
1 | | | | | | | | | | | | | | | | | | | |
Db 6294 TTTTGGATCCACAAGATTTCGTA 6315

```

Search completed: December 15, 2001, 03:20:02  
Job time: 7692 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 03:29:46 ; Search time 401.91 seconds  
(without alignments)  
46,929 Million cell updates/sec

Title: US-09-380-826a-7

Perfect score: 22

Sequence: 1 tgttgatcacagaattgata 22

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: N.Geneseq\_1101.\*  
2: /SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq/NA1986.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	100.0	22	19	AAV58901
2	22	100.0	22	19	AAV58897
3	22	100.0	1477	19	AAV58896
4	18.8	85.5	1012	21	AAV59365
5	18.8	85.5	1069	21	AAV59365
6	17.8	80.9	2084	21	AAV79696
7	17.2	78.2	447	21	AAV18114
8	17.2	78.2	1290	21	AAV42365
9	17.2	78.2	1413	21	AAV46258
10	17.2	78.2	2038	22	AAV22799
11	16.8	76.4	1764	22	AAH15352

C	12	16.8	76.4	2582	18	AAT90543	Maize male fertill
C	13	16.4	74.5	2444	21	AAC56059	Eucalyptus grandis
C	14	16.2	73.6	335	22	AAH37702	Human colon cancer
C	15	16.2	73.6	951	20	AAV84598	Human secreted pro
C	16	16.2	73.6	1250	21	AAV59054	Human secreted pro
C	17	16.2	73.6	1291	19	AAV59636	Human secreted pro
C	18	16.2	73.6	1376	22	AAH33007	Human colon cancer
C	19	16.2	73.6	1450	21	AAC49931	Arabidopsis thalia
C	20	16.2	73.6	1452	21	AAC39657	Arabidopsis thalia
C	21	16.2	73.6	1473	20	AAV61757	B. burgdorferi ant
C	22	16.2	73.6	2000	17	AAT71111	Acetyl-CoA-carboxy
C	23	16.2	73.6	2001	17	AAT39904	Maize acetyl CoA c
C	24	16.2	73.6	2001	21	AAZ49816	ECORI fragment of
C	25	16.2	73.6	3822	18	AAT64683	M. leprae gyrA pre
C	26	16.2	73.6	4346	14	AAO42933	A3 maize ACCase cd
C	27	16.2	73.6	5400	21	AAH37732	Potato subclone pg
C	28	16.2	73.6	7470	17	AAT39905	Maize acetyl CoA c
C	29	16.2	73.6	7470	19	AAV29317	Maize ACCase enzym
C	30	16.2	73.6	7470	21	AAZ49820	Maize acetyl CoA c
C	31	16.2	73.6	10811	19	AAV59091	Potato pollen cell
C	32	15.8	71.8	467	22	AAH32060	Human olfactory re
C	33	15.8	71.8	670	21	AAC09187	Human secreted pro
C	34	15.8	71.8	736	21	AAC10808	Human secreted pro
C	35	15.8	71.8	740	21	AAC51908	Arabidopsis thalia
C	36	15.8	71.8	843	20	AAZ15467	Human gene express
C	37	15.8	71.8	961	19	AAH14095	H. pylori GHP0 131
C	38	15.8	71.8	4146	21	AAZ51556	Human hypoxia resp
C	39	15.6	70.9	297	20	AAV88753	EST clone HK650.
C	40	15.6	70.9	342	22	AAH32175	Human olfactory re
C	41	15.6	70.9	452	22	AAH31651	S. epidermidis ope
C	42	15.6	70.9	635	22	AAH31988	Human olfactory re
C	43	15.6	70.9	795	20	AAV72025	Adenovirus PAC7SG2
C	44	15.6	70.9	834	20	AAV72026	Adenovirus SCAR-RG
C	45	15.6	70.9	1075	18	AAT72715	C. elegans inhibit

## ALIGNMENTS

RESULT 1	AAV58901	standard; DNA; 22 BP.
ID	AAV58901	
XX	AAV58901	
AC	AAV58901	
XX	20-JAN-1999	(first entry)
DT	XX	
XX	XX	
DE	XX	Leptospira rRNA gene nucleotide sequence.
XX	XX	Infection; pathogenic Leptospira; protective immunity; therapy;
KW	XX	diagnosis; ss.
KW	XX	
OS	XX	Leptospira sp.
XX	XX	
PN	XX	W09840099-A1.
XX	XX	
PD	XX	17-SEP-1998.
XX	XX	
PF	XX	06-MAR-1998; 98WO-AU00145.
XX	XX	
PR	XX	07-MAR-1997; 97AU-0005494.
XX	XX	
PA	XX	(AGRI-) AGRIC VICTORIA SERVICES PTY LTD.
PA	XX	(PIGR-) PIC RES & DEV CORP.
PI	XX	Chappel RJ;
XX	XX	
DR	XX	WPI; 1998-520791/44.
XX	XX	
PT	XX	New isolated pathogenic Leptospira bacterium - useful for, e.g
PT	XX	developing products for conferring protective immunity, and for
PT	XX	prophylactic or therapeutic treatment
XX	XX	

```
PS Claim 15; Page 72; 94pp; English.
XX
CC This sequence represents a Leptospira DNA sequence isolated from the
CC pathogenic Leptospira (LS) bacterium of the invention. The bacterium
CC belongs to serogroup Hurstbridge or serovar hurstbridge or the species
CC L. fainel. The LS bacteria can be used for conferring protective
CC immunity against pathogenic LS bacteria in humans or animals. The
CC bacteria can also be used for prophylactic or therapeutic treatment of LS
CC infections. The DNAs and antibodies may also be used for detection and
CC diagnosis of past or present LS infection.
XX
SQ Sequence 22 BP; 7 A; 2 C; 5 G; 8 T; 0 other;

Query Match      100.0%; Score 22; DB 19; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tgttgatcacagaattgata 22
   |||||||||||||||||||
Db 1 tgttgatcacagaattgata 22

RESULT 2
AAV58897
ID AAV58897 standard; DNA; 22 BP.
XX
AC AAV58897;
XX
DT 20-JAN-1999 (first entry)
XX
DE L. fainel nucleotide sequence.
XX
KW Infection; pathogenic Leptospira; protective immunity; therapy;
XX diagnosis; ss.
XX
KM Leptospira fainel.
XX
OS
XX
PN WO9840099-A1.
XX
PD 17-SEP-1998.
XX
PF 06-MAR-1998; 98WO-AU00145.
XX
PR 07-MAR-1997; 97AU-0005494.
XX
PA (AGRI-) AGRIC VICTORIA SERVICES PTY LTD.
XX (PIGR-) PIG RES & DEV CORP.
XX
PI Chappel RJ;
XX
DR WPI; 1998-520791/44.
XX
PT New isolated pathogenic Leptospira bacterium - useful for, e.g
PT developing products for conferring protective immunity, and for
PT prophylactic or therapeutic treatment
XX
PS Claim 15; Page 70; 94pp; English.
XX
CC This sequence represents a Leptospira DNA sequence isolated from the
CC pathogenic Leptospira (LS) bacterium of the invention. The bacterium
CC belongs to serogroup Hurstbridge or serovar hurstbridge or the species
CC L. fainel. The LS bacteria can be used for conferring protective
CC immunity against pathogenic LS bacteria in humans or animals. The
CC bacteria can also be used for prophylactic or therapeutic treatment of LS
CC infections. The DNAs and antibodies may also be used for detection and
CC diagnosis of past or present LS infection.
XX
SQ Sequence 22 BP; 7 A; 2 C; 5 G; 8 T; 0 other;

Query Match      100.0%; Score 22; DB 19; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 22; Conservative 100.0%; Mismatches 0; Indels 0; Gaps 0;

QY 1 tgttgatcacagaattgata 22
   |||||||||||||||||||
Db 1 tgttgatcacagaattgata 22

RESULT 3
AAV58896
ID AAV58896 standard; DNA; 1477 BP.
XX
AC AAV58896;
XX
DT 20-JAN-1999 (first entry)
XX
DE L. fainel nucleotide sequence.
XX
KW Infection; pathogenic Leptospira; protective immunity; therapy;
XX diagnosis; ss.
XX
KM Leptospira fainel.
XX
OS
XX
PN WO9840099-A1.
XX
PD 17-SEP-1998.
XX
PF 06-MAR-1998; 98WO-AU00145.
XX
PR 07-MAR-1997; 97AU-0005494.
XX
PA (AGRI-) AGRIC VICTORIA SERVICES PTY LTD.
XX (PIGR-) PIG RES & DEV CORP.
XX
PI Chappel RJ;
XX
DR WPI; 1998-520791/44.
XX
PT New isolated pathogenic Leptospira bacterium - useful for, e.g
PT developing products for conferring protective immunity, and for
PT prophylactic or therapeutic treatment
XX
PS Claim 15; Page 69-70; 94pp; English.
XX
CC This sequence represents a Leptospira DNA sequence isolated from the
CC pathogenic Leptospira (LS) bacterium of the invention. The bacterium
CC belongs to serogroup Hurstbridge or serovar hurstbridge or the species
CC L. fainel. The LS bacteria can be used for conferring protective
CC immunity against pathogenic LS bacteria in humans or animals. The
CC bacteria can also be used for prophylactic or therapeutic treatment of LS
CC infections. The DNAs and antibodies may also be used for detection and
CC diagnosis of past or present LS infection.
XX
SQ Sequence 1477 BP; 390 A; 334 C; 439 G; 314 T; 0 other;

Query Match      100.0%; Score 22; DB 19; Length 1477;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tgttgatcacagaattgata 22
   |||||||||||||||||||
Db 154 tgttgatcacagaattgata 175

RESULT 4
AAA99365
ID AAA99365 standard; DNA; 1012 BP.
XX
AC AAA99365;
XX
DT 22-JAN-2001 (first entry)
XX
DE Plant PrG1 promoter related gene sequence.
```

XX Plant promoter; PRAG1; reproductive tissue; transgenic plant; cereal; ds.  
KW Arabidopsis thaliana.  
OS WO20005172-A1.  
PN 21-SEP-2000.  
XX 17-MAR-2000; 2000WO-N200031.  
XX 17-MAR-1999; 99NZ-0334715.  
XX (CART-) CARTER HOLT HARVEY LTD.  
PA (TASMAN BIOTECHNOLOGY LTD.  
PA (UNMKT ) UNIV MICHIGAN TECHNOLOGICAL.  
XX Podilla GK, Liu J, Karnosky DF;  
PI WPI: 2000-594442/56.  
DR P-PSDB; AAB26796.  
XX Novel plant reproductive tissue promoter, useful to produce plants  
PT which have a diminished reproductive capacity or which are sterile  
XX Claim 15; Page 42-43; 51pp; English.  
XX This invention relates to a novel plant promoter gene. The promoter is  
CC located in plant reproductive tissue, and the invention includes  
CC transgenic plants containing the promoter. The promoter can be used to  
CC produce plants which have a diminished reproductive capacity or which are  
CC sterile. The constructs can also be used to transform agronomically  
CC important plants in which modulation of reproductive capacity  
CC (particularly the timing and abundance of flowering) is desirable,  
CC e.g. cereals, rice, maize, wheat, barley, oats, rye, soybean and canola.  
CC The present sequence represents DNA encoding the plant reproductive  
CC promoter (PRAG1) of the invention.  
XX Sequence 1012 BP; 287 A; 197 C; 217 G; 311 T; 0 other;  
SQ

Query Match 85.5%; Score 18.8; DB 21; Length 1012;  
Best Local Similarity 90.9%; Pred. No. 5.8;  
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Oy 1 ttttgatcacaaagtattgata 22  
||||| ||||||||| 4  
Db 672 ttttgatcacaaagtattgaca 693

RESULT 5  
AAC37325  
ID AAC37325 standard; DNA; 1069 BP.  
XX AAC37325;  
AC 17-OCT-2000 (first entry)  
DT 17-OCT-2000 (first entry)  
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 16971.  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 16971.  
XX Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway;  
KW metabolic pathway; promoter; termination sequence; ss.  
XX Arabidopsis thaliana.  
XX Arabidopsis thaliana.  
PN EP1033405-A2.  
PD 06-SEP-2000.  
XX 25-FEB-2000; 2000EP-0301439.  
PF 25-FEB-1999; 99US-0121825.  
XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.

```
PR 14-JUL-1999; 99US-0143624.*
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145324.
PR 23-JUL-1999; 99US-0145276.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148365.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156506.
PR 04-OCT-1999; 99US-0157117.
```

```
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161993.
PR 28-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.

Query Match 85.5%; Score 18.8; DB 21; Length 1069;
Best Local Similarity 90.5%; Pred. No. 5.8;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 tgttgatcacaaagtgtgata 22
Db 690 tgttgatcacaaagtgtgata 711

RESULT 6
AAA79696/c
ID AAA79696 standard; cDNA; 2084 BP.
XX
AC AAA79696;
XX
DT 27-NOV-2000 (first entry)
XX
DE
XX Eucalyptus grandis cell signalling involved polynucleotide SEQ ID NO:848.
XX
KW Eucalyptus grandis; Pinus radiata; Monterey pine; plant; modification;
KW plant cell signalling; modulation; transgenic plant; pathogen; growth;
KW environmental change; development; cell proliferation; differentiation;
KW elongation; survival; disease resistance; nutrient metabolism; ss.
XX
OS Eucalyptus grandis.
XX
PN WO200042171-A1.
XX
PD 20-JUL-2000.
XX
PF 11-JAN-2000; 2000MO-US00724.
XX
PR 12-JAN-1999; 99US-0228986.
PR 01-NOV-1999; 99US-0162866.
XX
PA (GENE-) GENESIS RES & DEV CORP LTD.
XX
PI Strabala TJ, Nieuwenhuizen NJ;
```

DR WPI; 2000-476052/41.  
XX Isolated polynucleotide encoding a polypeptide involved in cell  
PT signaling used for generating transgenic plants with modified responses  
PT to external signals -  
PS Claim 1; Page 408-409; 527pp; English.  
XX  
CC AAA79263 to AAA79736 and AAB25100 to AAB25570 represent polynucleotide  
CC and protein sequences isolated from eucalyptus (*Eucalyptus grandis*) or  
CC pine (*Pinus radiata* also known as Monterey pine). The protein sequences  
CC are involved in cell signaling. The polynucleotide and protein  
CC sequences can be used to modify the response of plant cells to external  
CC signals e.g. environmental changes or pathogens during the growth and  
CC development of a plant. They can be used to modify cell proliferation,  
CC differentiation, elongation and survival, resistance to disease and  
CC nutrient metabolism. Examples of modifications which can be produced are  
CC altered fruit ripening and senescence of leaves and flowers e.g. to  
CC delay senescence and prolong the life of cut flowers or enhance  
CC senescence of reproductive organs to engineer sterile plants. Other  
CC modifications can be used to delay senescence in selected cell types or  
CC organs providing fruit and vegetables which have a longer shelf life  
CC between harvest and consumption, or to decrease branching frequency in  
CC forest tree species giving long stretches of valuable knot-free clear  
CC wood which can be used in solid timber furniture and veneers.  
XX  
SQ Sequence 2084 BP; 502 A; 434 C; 566 G; 582 T; 0 other;

Query Match 80.9%; Score 17.8; DB 21; Length 2084;  
Best Local Similarity 90.5%; Pred. No. 19;  
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 tgttgatcacagaattgat 21  
||||| ||||| ||||| |||||  
DB 1347 tcttgagcagcaagtttgat 1327

RESULT 7  
AAC18114  
ID AAC18114 standard; cDNA; 447 BP.  
XX  
AC AAC18114;  
XX  
DT 06-OCT-2000 (first entry)  
XX  
DE Human secreted protein 5' EST, SEQ ID NO: 22189.  
XX  
KM Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
KM gene therapy; chromosome mapping; ss.  
XX  
OS Homo sapiens.  
PN EP1033401-A2.  
PD 06-SEP-2000.  
XX  
PF 21-FEB-2000; 2000EP-0200610.  
XX  
PR 26-FEB-1999; 9905-0122487.  
XX  
PA (GEST ) GENSET.  
XX  
PI Dumas Mline Edwards J, Duclert A, Giordano J;  
DR WPI; 2000-500381/45.  
XX  
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -  
XX  
PS Claim 1; SEQ ID 22189; 71pp + CD-ROM; English.  
XX

CC The present sequence is one of a large number of 5' ESTs derived from  
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively  
CC identified within the present sequence. The 5' ESTs were prepared from  
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST  
CC sequences usually correspond mainly to the 3' untranslated region (UTR)  
CC of the mRNA because they are often obtained from oligo-dT primed cDNA  
CC libraries. Such ESTs are not well suited for isolating cDNA sequences  
CC derived from the 5' ends of mRNAs and even in those cases where longer  
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.  
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be  
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used  
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.  
CC They are used to obtain upstream regulatory sequences and to design  
XX expression and secretion vectors.  
XX  
SQ Sequence 447 BP; 187 A; 57 C; 66 G; 128 T; 9 other;

Query Match 78.2%; Score 17.2; DB 21; Length 447;  
Best Local Similarity 86.4%; Pred. No. 32;  
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 tgttgatcacagaattgata 22  
||||| ||||| || || ||  
DB 240 tgttgatcacacaacttaata 261

RESULT 8  
AAC42365/C  
ID AAC42365 standard; DNA; 1290 BP.  
XX  
AC AAC42365;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 35275.  
XX  
KM Hybridisation assay; genetic mapping; gene expression control;  
KM protein identification; signal transduction pathway;  
KM metabolic pathway; promoter; termination sequence; ss.  
XX  
OS Arabidopsis thaliana.  
PN EP1033405-A2.  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 9905-0121825.  
PR 05-MAR-1999; 9905-0123180.  
PR 09-MAR-1999; 9905-0123548.  
PR 23-MAR-1999; 9905-0125788.  
PR 25-MAR-1999; 9905-0126264.  
PR 29-MAR-1999; 9905-0126785.  
PR 01-APR-1999; 9905-0127462.  
PR 06-APR-1999; 9905-0128234.  
PR 08-APR-1999; 9905-0128714.  
PR 16-APR-1999; 9905-0129845.  
PR 19-APR-1999; 9905-0130077.  
PR 21-APR-1999; 9905-0130449.  
PR 23-APR-1999; 9905-0130510.  
PR 23-APR-1999; 9905-0130891.  
PR 28-APR-1999; 9905-0131449.  
PR 30-APR-1999; 9905-0132048.  
PR 30-APR-1999; 9905-0132407.  
PR 04-MAY-1999; 9905-0132484.  
PR 05-MAY-1999; 9905-0132485.  
PR 06-MAY-1999; 9905-0132486.  
PR 07-MAY-1999; 9905-0132487.  
PR 11-MAY-1999; 9905-0132863.  
PR 14-MAY-1999; 9905-0134256.  
PR 14-MAY-1999; 9905-0134218.

PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139464.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.



PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161921.  
PR 28-OCT-1999; 99US-0161922.  
PR 28-OCT-1999; 99US-0161923.  
PR 29-OCT-1999; 99US-0162142.  
  
Query Match 78.2%; Score 17.2; DB 21; Length 1290;  
Best Local Similarity 86.4%; Pred. No. 36;  
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
Oy 1 ttttgatcacagaatttcata 22  
|||||t|||||||t|t|  
Db 642 TGTTCGTACCAAGATTTCGTA 621  
  
RESULT 9  
AAC46258/c  
ID AAC46258 standard; DNA; 1413 BP.  
XX  
AC AAC46258;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 49483.  
XX  
KW Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway;  
KW metabolic pathway; promoter; termination sequence; ss.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
PD  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301433;  
XX  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0134256.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135153.

PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.

PR 02-AUG-1999; 99US-0146389.  
 PR 03-AUG-1999; 99US-0147038.  
 PR 04-AUG-1999; 99US-0147204.  
 PR 04-AUG-1999; 99US-0147302.  
 PR 05-AUG-1999; 99US-0147192.  
 PR 05-AUG-1999; 99US-0147260.  
 PR 06-AUG-1999; 99US-0147303.  
 PR 06-AUG-1999; 99US-0147416.  
 PR 09-AUG-1999; 99US-0147493.  
 PR 09-AUG-1999; 99US-0147935.  
 PR 10-AUG-1999; 99US-0148171.  
 PR 11-AUG-1999; 99US-0148319.  
 PR 12-AUG-1999; 99US-0148341.  
 PR 13-AUG-1999; 99US-0148565.  
 PR 13-AUG-1999; 99US-0148684.  
 PR 16-AUG-1999; 99US-0149368.  
 PR 17-AUG-1999; 99US-0149175.  
 PR 18-AUG-1999; 99US-0149426.  
 PR 20-AUG-1999; 99US-0149722.  
 PR 20-AUG-1999; 99US-0149723.  
 PR 20-AUG-1999; 99US-0149929.  
 PR 23-AUG-1999; 99US-0149902.  
 PR 23-AUG-1999; 99US-0149930.  
 PR 25-AUG-1999; 99US-0150566.  
 PR 26-AUG-1999; 99US-0150884.  
 PR 27-AUG-1999; 99US-0151085.  
 PR 27-AUG-1999; 99US-0151086.  
 PR 27-AUG-1999; 99US-0151086.  
 PR 30-AUG-1999; 99US-0151303.  
 PR 31-AUG-1999; 99US-0151438.  
 PR 01-SEP-1999; 99US-0151930.  
 PR 07-SEP-1999; 99US-0152363.  
 PR 10-SEP-1999; 99US-0153070.  
 PR 13-SEP-1999; 99US-0153758.  
 PR 15-SEP-1999; 99US-0154018.  
 PR 16-SEP-1999; 99US-0154039.  
 PR 20-SEP-1999; 99US-0154779.  
 PR 22-SEP-1999; 99US-0155139.  
 PR 23-SEP-1999; 99US-0155486.  
 PR 24-SEP-1999; 99US-0155659.  
 PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 13-OCT-1999; 99US-0159295.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 22-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.

PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.  
 Query Match 78.2%; Score 17.2; DB 21; Length 1413;  
 Best Local Similarity 86.4%; Pred. No. 36;  
 Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 Qy 1 lgttgatcacaagatttgata 22  
 Db 576 TGTGTACACAGATTGTGA 555  
 RESULT 10  
 AAS22799  
 ID AAS22799 standard; CDNA: 2038 BP.  
 AC AAS22799;  
 XX  
 XX  
 DT 24-OCT-2001 (first entry)  
 XX  
 DE Human cDNA encoding a novel human protein #365.  
 XX  
 KW Human; novel protein; se; Antianemic; osteopathic; antiinflammatory;  
 KW immunomodulatory; cytosolic; neuroprotective; vulnery; neotropic;  
 KW anticonvulsant; antiallergic; cerebroprotective; antifungal; antiviral;  
 KW antibacterial; antiallergic; dermatological; hemostatic; antistatic;  
 KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;  
 KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;  
 KW tissue regeneration; immune disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200155437-A2.  
 XX  
 PD 02-AUG-2001.  
 XX  
 XX 25-JAN-2001; 2001MO-US02623.  
 PF  
 XX 25-JAN-2000; 2000US-0491404.  
 XX  
 PA (HYSEQ-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Drmanac RT;  
 XX  
 DR WPI: 2001-451939/48.  
 XX  
 PT P-PSDB: AAU14494.  
 XX  
 PT Isolated polypeptides useful for treating anti-inflammatory diseases,  
 XX nervous system disorders, and for regenerating bone and cartilage -  
 PS Claim 1; Page 733-734; 894pp; English.  
 XX  
 CC The invention relates to polynucleotides encoding novel human  
 CC proteins or their active domains. The polypeptides, polynucleotides and  
 CC antibodies raised against the polypeptides are used in a method of  
 CC treatment of a mammal and prevention of disorders caused by the aberrant  
 CC protein expression or activity. The polypeptides can be used as  
 CC molecular weight markers, food supplements, and in antibody production.  
 CC The polypeptides are used to identify compounds which bind to the  
 CC polypeptides. Polynucleotides of the invention are used as probes and  
 CC primers, for sequencing, for chromosome or gene mapping, in the  
 CC production of recombinant proteins, and in generating anti-sense DNA or  
 CC RNA and in gene therapy. Polypeptides of the invention can be used to  
 CC target drugs to a tumour, in assays to determine biological activity, to  
 CC raise antibodies/elicite an immune response, to determine receptors or ligands.  
 CC Polypeptides of the invention may also be useful in treating platelet  
 CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,  
 CC ligament and/or nerve tissue, wound healing, treating burns, promoting  
 CC the proliferation, differentiation and survival of stem cells, as a  
 CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,  
 CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral



be derived from 231 via genetic rearrangements, deletions and/or insertions. 233 has subsequent deletions from 235 and is truncated, having only exons 3, 5 and 6. The M41-A and Zm41-A genes, antisense or ribozyme sequences can be used to produce transgenic plants with controlled male fertility. Male sterile plants are useful for hybrid seed production, particularly in plants where restoration of fertility is not needed, e.g. Brassicaceae, lettuce, spinach and onions.

Sequence 2582 BP; 670 A; 512 C; 537 G; 863 T; 0 other;

Query Match 76.4%; Score 16.8; DB 18; Length 2582;  
Best Local Similarity 90.0%; Pred. No. 60;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tgttgatcacagaattga 20  
||||| ||||| ||||| |||||  
Db 268 TGTGATCACAGATCTGA 249

## RESULT 13

AAC56059 standard; DNA: 2444 BP.  
AAC56059;

25-JAN-2001 (first entry)

Eucalyptus grandis transcription factor DNA sequence #190.

Plant; transcription factor; gene expression; eucalyptus; pine; acacia; poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor; basic helix-loop-helix zipper; homeotic; homeobox; MADS; homeodomain zipper; LIM domain; AP2; EREBs; zinc finger domain; type 2 Cys2His2; CCAAT box element; MYB; ss.

Eucalyptus grandis.

WO200053724-A2.

14-SEP-2000.  
09-MAR-2000; 2000MO-US06112.

11-MAR-1999; 99US-0266513.  
18-AUG-1999; 99US-0149483.

(GENE-) GENESIS RES & DEV CORP LTD.  
(FLET-) FLETCHER CHALLENGE FORESTS LTD.

Wood M, McGrath A, Shenk MA, Glenn M;

WPI: 2000-579369/54.

New isolated polynucleotide encoding a plant transcription factor for producing a plant e.g. a woody plant, preferably eucalyptus or pine, having modified gene expression or modified activity of a polypeptide

Claim 1; Pages 94-95; 747pp; English.

The present invention relates to novel plant transcription factors from Eucalyptus grandis or Pinus radiata. The present sequence is the coding sequence for one such transcription factor. The transcription factor may be used to produce a plant having modified gene expression such as a woody plant e.g. a eucalyptus, pine, acacia, poplar, sweetgum, teak, or mahogany species or to modify the activity of a polypeptide in a plant. The transcription factors of the present invention are members from the following families of regulatory proteins: bZIP, bZIP family of G-box binding factors, basic helix-loop-helix zipper, LIM domain, AP2 homeotic/homeodomain/homeobox/MADS, homeobox zipper, CCAAT box elements and EREBs, zinc finger domains of type 2 Cys2His2, CCAAT box elements

CC and MYB.  
XX Sequence 2444 BP; 643 A; 554 C; 640 G; 607 T; 0 other;  
SQ

Query Match 74.5%; Score 16.4; DB 21; Length 2444;  
Best Local Similarity 94.4%; Pred. No. 94;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 tgttgatcacagaattc 18  
||||| ||||| ||||| |||||  
Db 1451 tgttgatcacagaattc 1468

RESULT 14  
AAH33702/C  
ID AAH33702 standard; CDNA: 335 BP.

AAH33702;

03-SEP-2001 (first entry)

Human colon cancer antigen encoding cDNA SEQ ID NO:758.

Human; colon cancer; colon cancer antigen; diagnosis; detection; colorectal carcinoma; ss.

Homo sapiens.

WO200122920-A2.

05-APR-2001.

28-SEP-2000; 2000MO-US26524.

29-SEP-1999; 99US-0157137.

03-NOV-1999; 99US-0163280.

(HUMA-) HUMAN GENOME SCI INC.

Ruben SM, Barash SC, Birse CE, Rosen CA;

WPI: 2001-235357/24.

P-PSDB; AAG74271.

Nucleic acids encoding 427 human colon cancer-associated polypeptides, useful for preventing, diagnosing and/or treating colorectal cancers -

Claim 1; Page 2738; 9803pp; English.  
AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon cancer-associated nucleic acid molecules (N) and proteins (P), where the proteins are collectively known as colon cancer antigens. The colon cancer antigens have cytostatic activity and can be used in gene therapy and vaccine production. N and P may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate P expression. For example, N and P may be used to treat disorders CC associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of P by expressing

inactive proteins or to supplement the patients own production of P. Additionally, N may be used to produce the colon cancer-associated P, by inserting the nucleic acids into a host cell and culturing the cell to express the proteins. N and P can be used in the prevention, diagnosis and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204 and AAG77789 represent sequences used in the exemplification of the present invention.

N.B. Pages 666 to 682 and page 7053 of the sequence listing were missing at time of publication, meaning no sequences are present for SEQ ID NO:1027 to 1052, 7921 and 7922.

Sequence 335 BP; 95 A; 71 C; 60 G; 105 T; 4 other;



CC restenosis, prostate diseases, obesity, disorders involving osteoclasts  
 CC such as osteoporosis, arthritis or malignancies, diseases of testes, the  
 CC lung or thymus, digestive/endocrine disorders, infections and AIDS. The  
 CC polypeptides are also useful for identifying their binding partners.  
 CC The present sequence represents a gene encoding a human secreted protein  
 CC (see descriptor line for gene number and clone identification).

XX  
 SQ Sequence 951 BP; 296 A; 134 C; 192 G; 329 T; 0 other;

Query Match 73.6%; Score 16.2; DB 20; Length 951;  
 Best Local Similarity 85.7%; Pred. No. 1.1e+02;  
 Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 tgltgatcacagaattgat 21  
 ||||| || ||||| |||||  
 Db 524 TGTGGTTCTCAAGATCTGAT 504

Search completed: December 15, 2001, 03:29:47  
 Job time: 7887 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 03:22:49 ; Search time 172.39 Seconds  
(without alignments)  
28.903 Million cell updates/sec

Title: US-09-380-826A-7

Perfect score: 22

Sequence: 1 ttttgatcacaaagattgata 22

Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 351203 segs, 11323899 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_NA:\*

1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PCUTS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/Backfillseq1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB* ID	Description
c 1	16.8	76.4	2582	4 US-09-102-528-31	Sequence 31, Appl
c 2	16.2	73.6	2000	1 US-08-014-326-1	Sequence 1, Appl
c 3	16.2	73.6	2001	3 US-08-417-089-1	Sequence 1, Appl
c 4	16.2	73.6	2001	3 US-08-695-651-1	Sequence 1, Appl
c 5	16.2	73.6	2001	4 US-08-930-285-1	Sequence 1, Appl
c 6	16.2	73.6	2001	4 US-08-695-421-1	Sequence 1, Appl
c 7	16.2	73.6	4345	2 US-08-244-537-1	Sequence 1, Appl
c 8	16.2	73.6	7470	3 US-08-417-089-5	Sequence 1, Appl
c 9	16.2	73.6	7470	3 US-08-695-651-5	Sequence 1, Appl
c 10	16.2	73.6	7470	4 US-08-930-285-5	Sequence 5, Appl
c 11	16.2	73.6	7470	4 US-08-695-421-5	Sequence 5, Appl
c 12	15.6	70.9	296	3 US-08-602-145-14	Sequence 14, Appl
c 13	15.6	70.9	1095	4 US-08-928-383B-3	Sequence 3, Appl
c 14	15.6	70.9	1096	1 US-08-684-862-8	Sequence 8, Appl
c 15	15.6	70.9	1584	4 US-08-928-383B-1	Sequence 1, Appl
c 16	15.6	70.9	1989	2 US-08-792-055-1	Sequence 1, Appl
c 17	15.6	70.9	2434	4 US-09-272-486-1	Sequence 1, Appl
c 18	15.6	70.9	2770	4 US-09-008-697A-13	Sequence 13, Appl
c 19	15.6	70.9	7739	4 US-09-193-966-1	Sequence 1, Appl
c 20	15.6	70.9	7745	1 US-08-299-675-1	Sequence 1, Appl
c 21	15.6	70.9	7745	2 US-08-485-241-1	Sequence 1, Appl
c 22	15.6	70.9	7745	2 US-08-874-162-1	Sequence 1, Appl
c 23	15.2	69.1	1016	1 US-08-399-986B-3	Sequence 3, Appl
c 24	15.2	69.1	1016	1 US-08-493-754A-3	Sequence 3, Appl
c 25	15.2	69.1	2182	1 US-08-399-986B-1	Sequence 1, Appl
c 26	15.2	69.1	2182	1 US-08-493-754A-1	Sequence 1, Appl
c 27	15.2	69.1	2598	3 US-08-745-892-20	Sequence 20, Appl

c 28	15.2	69.1	5420	6 5256642-3	Patent No. 5256642
c 29	15.2	69.1	5420	6 5472939-3	Patent No. 5472939
c 30	15.2	69.1	5872	4 US-09-102-528-32	Patent No. 5256642
c 31	15.2	69.1	6951	6 5256642-1	Patent No. 5472939
c 32	15.2	69.1	6951	6 5472939-1	Patent No. 5472939
c 33	14.8	67.3	337	4 US-08-991-789A-13	Sequence 13, Appl
c 34	14.8	67.3	1855	1 US-08-961-083-71	Sequence 71, Appl
c 35	14.8	67.3	3159	1 US-08-119-361-4	Sequence 4, Appl
c 36	14.8	67.3	3159	3 US-08-336-308A-3	Sequence 3, Appl
c 37	14.8	67.3	3159	3 US-08-822-324-3	Sequence 3, Appl
c 38	14.8	67.3	3159	4 US-09-490-931-3	Sequence 9, Appl
c 39	14.8	67.3	7266	3 US-08-336-308A-9	Sequence 9, Appl
c 40	14.8	67.3	7266	3 US-08-822-324-5	Sequence 5, Appl
c 41	14.8	67.3	7266	4 US-09-490-931-9	Sequence 9, Appl
c 42	14.8	67.3	8640	1 US-08-570-311-28	Sequence 28, Appl
c 43	14.8	67.3	9919	3 US-08-880-179-1	Sequence 1, Appl
c 44	14.6	66.4	293	3 US-08-866-340-13	Sequence 13, Appl
c 45	14.6	66.4	293	4 US-09-103-875-17	Sequence 17, Appl

## ALIGNMENTS

```
RESULT 1
US-09-102-528-31/c
; Sequence 31, Application US/09102528
; Patent No. 6207883
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: DNA Sequences
; NUMBER OF SEQUENCES: 32
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/102,528
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/GB96/03191
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2582 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-102-528-31

Query Match 76.4%; Score 16.8; DB 4; Length 2582;
Best Local Similarity 90.0%; Pred. No. 14;
Matches 18; Conservative 2; Indels 0; Gaps 0;

QY 1 ttttgatcacaaagattga 20
Db 268 ttttgatcacaaagattga 249

RESULT 2
US-08-014-326-1
; Sequence 1, Application US/08014326
; Patent No. 5498544
; GENERAL INFORMATION:
; APPLICANT: Gengenbach, Burle G.
; APPLICANT: Somers, David A.
; APPLICANT: Wyse, Donald L.
; APPLICANT: Gronwald, John W.
; APPLICANT: Egli, Margaret A.
; APPLICANT: Lutz, Shlela M.
; TITLE OF INVENTION: Method and An Acetyl CoA Carboxylase
```

```
; TITLE OF INVENTION: Alteration in Oil Content of Plants
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESSES:
; ADDRESS: Merchant & Gould
; STREET: 3100 No. 5498544west Center.
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/014,326
; FILING DATE: 05-FEB-1993
; CLASSIFICATION: 800
; PRIORITY INFORMATION:
; APPLICATION NUMBER: US 07/917,462
; FILING DATE: 21-JUL-1992
; PRIORITY INFORMATION:
; APPLICATION NUMBER: US 07/538,674
; FILING DATE: 18-JUN-1990
; PRIORITY INFORMATION:
; APPLICATION NUMBER: US 07/269,584
; FILING DATE: 10-NOV-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Kowalchuk, Katherine M.
; REGISTRATION NUMBER: 36,848
; REFERENCE/DOCKET NUMBER: 600.258-US01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-332-5300
; TELEFAX: 612-332-9081
; INFORMATION FOR SEQ. ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2000 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: 2 kb fragment of lambda clone #15-14
;
US-08-014-326-1
;
Query Match 73.6%; Score 16.2; DB 1; Length 2000;
Best Local Similarity 85.7%; Pred. No. 27;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
;
QY 2 gtcgatacacaagattgata 22
Db 42 GTTGATGACACAGAGTGTGA 62
;
RESULT 3
US-08-417-089-1
; Sequence 1, Application US/08417089
; Patent No. 6069298
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: METHODS AND AN ACETYL COA CARBOXYLASE GENE
; TITLE OF INVENTION: FOR CONFERRING HERBICIDE TOLERANCE AND AN ALTERATION IN
; NUMBER OF SEQUENCES: 11
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (RPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/417,089
; INFORMATION FOR SEQ. ID NO: 1:
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2001 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
;
US-08-417-089-1
;
Query Match 73.6%; Score 16.2; DB 3; Length 2001;
Best Local Similarity 85.7%; Pred. No. 27;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
;
QY 2 gtcgatacacaagattgata 22
Db 42 GTTGATGACACAGAGTGTGA 62
;
RESULT 4
US-08-695-651-1
; Sequence 1, Application US/08695651
; Patent No. 6146867
; GENERAL INFORMATION:
; APPLICANT: Gengenbach, B. G.
; APPLICANT: Somers, D. A.
; APPLICANT: Wyse, D. L.
; APPLICANT: Gronwald, J. W.
; APPLICANT: Egli, M. A.
; APPLICANT: Lutz, S. M.
; TITLE OF INVENTION: METHOD FOR PRODUCING PLANT ACETYL COA
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P. A.
; STREET: P.O. Box 2938.
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/695,651
; FILING DATE: 12-AUG-1996
; CLASSIFICATION: 435
; PRIORITY INFORMATION:
; APPLICATION NUMBER: 08/417089
; FILING DATE: 05-APR-1995
; APPLICATION NUMBER: 08/014326
; FILING DATE: 05-FEB-1993
; APPLICATION NUMBER: 07/917462
; FILING DATE: 21-JUL-1992
; APPLICATION NUMBER: 07/538674
; FILING DATE: 18-JUN-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Woessner, Warren D
; REGISTRATION NUMBER: 30,440
; REFERENCE/DOCKET NUMBER: 600.318US3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-373-6900
; TELEFAX: 612-339-3061
; TELEX:
; INFORMATION FOR SEQ. ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2001 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
;
US-08-695-651-1
```



Query Match 73.6%; Score 16.2; DB 3; Length 2001;  
Best Local Similarity 85.7%; Pred. No. 27;  
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 gttgatacacaagattgata 22  
||||| ||||| ||| ||  
DB 42 GTTGATGACACAGACTGTGTTA 62

RESULT 5  
US-08-930-285-1  
Sequence 1, Application US/08930285  
Patent No. 6222099  
GENERAL INFORMATION:  
APPLICANT: Regents of the University of Minnesota, et al.  
TITLE OF INVENTION: TRANSGENIC PLANTS EXPRESSING ACETYL COA CARBOXYLASE GE  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P. A.  
STREET: P. O. Box 2938  
CITY: Minneapolis  
STATE: MN  
COUNTRY: USA  
ZIP: 55402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/930.285  
FILING DATE: 13-APR-1998  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/04625  
FILING DATE: 04-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Embretson, Janet E.  
REGISTRATION NUMBER: 39,665  
REFERENCE/DOCKET NUMBER: 600.318054  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 612-339-0331  
TELEFAX: 612-339-3061  
TELEX:  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2001 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
US-08-930-285-1

Query Match 73.6%; Score 16.2; DB 4; Length 2001;  
Best Local Similarity 85.7%; Pred. No. 27;  
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 gttgatacacaagattgata 22  
||||| ||||| ||| ||  
DB 42 GTTGATGACACAGACTGTGTTA 62

RESULT 6  
US-08-695-421-1  
Sequence 1, Application US/08695421  
Patent No. 6268550  
GENERAL INFORMATION:

APPLICANT: Gengenbach, B. G.  
APPLICANT: Somers, D. A.  
APPLICANT: Wyse, D. L.  
APPLICANT: Gronwald, J. W.  
APPLICANT: Egli, M. A.  
APPLICANT: Lutz, S. M.  
TITLE OF INVENTION: METHODS AND AN ACETYL COA CARBOXYLASE GENE  
TITLE OF INVENTION: FOR CONFERRING HERBICIDE TOLERANCE AND AN ALTERATION IN OIL  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P. A.  
STREET: P. O. Box 2938  
CITY: Minneapolis  
STATE: MN  
COUNTRY: USA  
ZIP: 55402

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/695.421  
FILING DATE: 23-AUG-1996  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/417089  
FILING DATE: 05-APR-1995  
APPLICATION NUMBER: 08/014326  
FILING DATE: 05-FEB-1993  
APPLICATION NUMBER: 07/917462  
FILING DATE: 21-JUL-1992  
APPLICATION NUMBER: 07/538674  
FILING DATE: 18-JUN-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Woessner, Warren D.  
REGISTRATION NUMBER: 30,440  
REFERENCE/DOCKET NUMBER: 600.318052  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 612-373-6900  
TELEFAX: 612-339-3061  
TELEX:

INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2001 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-08-695-421-1

Query Match 73.6%; Score 16.2; DB 4; Length 2001;  
Best Local Similarity 85.7%; Pred. No. 27;  
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 gttgatacacaagattgata 22  
||||| ||||| ||| ||  
DB 42 GTTGATGACACAGACTGTGTTA 62

RESULT 7  
US-08-244-537-1  
Sequence 1, Application US/08244537  
Patent No. 5854420  
GENERAL INFORMATION:  
APPLICANT: ASHTON, ANTHONY R.  
APPLICANT: JENKINS, COLIN L.D.  
APPLICANT: WHITEFIELD, PAUL R.  
TITLE OF INVENTION: MAIZE ACETYL COA CARBOXYLASE ENCODING  
TITLE OF INVENTION: DNA CLONES  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:

ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.  
STREET: 1100 New York Avenue, N.W.  
CITY: Washington  
STATE: D. C.  
COUNTRY: U.S.A.  
ZIP: 20005-3918  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/244,537  
FILING DATE: 18-AUG-1994  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9125330.2  
FILING DATE: 28-NOV-1991  
PRIOR APPLICATION DATA: PCT/GB92/02205  
FILING DATE: 27-NOV-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: KOKULIS, PAUL N.  
REGISTRATION NUMBER: 16,773  
REFERENCE/DOCKET NUMBER: 203094/SEE 36663/UST  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 861-3000  
TELEFAX: (202) 822-0944  
TELEX: 6714627 CUSH  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4345 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: both  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-244-537-1

Query Match 73.6% Score 16.2; DB 2; Length 4345;  
Best Local Similarity 85.7% Pred. No. 30;  
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 gtgagtcacagaattgata 22  
||||| ||||| ||| ||

Db 902 GTTGATGACAGAAGTTGTTA 922

RESULT 8  
US-08-417-089-5  
Sequence 5, Application US/08417089  
Patent No. 6069298  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: METHODS AND AN ACETYL COA CARBOXYLASE GENE  
TITLE OF INVENTION: FOR CONFERRING HERBICIDE TOLERANCE AND AN ALTERATION IN  
TITLE OF INVENTION: OIL CONTENT OF PLANTS  
NUMBER OF SEQUENCES: 11  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/417,089  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7470 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-417-089-5

Query Match 73.6% Score 16.2; DB 3; Length 7470;  
Best Local Similarity 85.7% Pred. No. 32;  
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 gtgagtcacagaattgata 22  
||||| ||||| ||| ||

Db 3972 GTTGATGACAGAAGTTGTTA 3992

RESULT 9  
US-08-695-651-5  
Sequence 5, Application US/08695651  
Patent No. 6146867  
GENERAL INFORMATION:  
APPLICANT: Gengenbach, B. G.  
APPLICANT: Somers, D. A.  
APPLICANT: Wyse, D. L.  
APPLICANT: Gronwald, J. W.  
APPLICANT: Egli, M. A.  
APPLICANT: Lutz, S. M.  
TITLE OF INVENTION: METHOD FOR PRODUCING PLANT ACETYL COA  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.  
STREET: P.O. Box 2938  
CITY: Minneapolis  
STATE: MN  
COUNTRY: USA  
ZIP: 55402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/695,651  
FILING DATE: 12-AUG-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/417089  
FILING DATE: 05-APR-1995  
APPLICATION NUMBER: 08/014326  
FILING DATE: 05-FEB-1993  
APPLICATION NUMBER: 07/917462  
FILING DATE: 21-JUL-1992  
APPLICATION NUMBER: 07/538674  
FILING DATE: 18-JUN-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Woessner, Warren D  
REGISTRATION NUMBER: 30,440  
REFERENCE/DOCKET NUMBER: 600.318053  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 612-373-6900  
TELEFAX: 612-339-3061  
TELEX:  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7470 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-08-695-651-5

Query Match 73.6% Score 16.2; DB 3; Length 7470;  
Best Local Similarity 85.7% Pred. No. 32;  
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 gtgagtcacagaattgata 22  
||||| ||||| ||| ||

DB 3972 GTTGATGACAGAGTTGTTA 3992

## RESULT 10

US-08-930-285-5  
; Sequence 5, Application US/08930285  
; Patent No. 6222099  
; GENERAL INFORMATION:  
; APPLICANT: Regents of the University of Minnesota, et al.  
; TITLE OF INVENTION: TRANSGENIC PLANTS EXPRESSING ACETYL COA CARBOXYLASE GE  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P. A.  
; STREET: P. O. Box 2938  
; CITY: Minneapolis  
; STATE: MN  
; COUNTRY: USA  
; ZIP: 55402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/930,285  
; FILING DATE: 13-APR-1998  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US96/04625  
; FILING DATE: 04-APR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Embretson, Janet E.  
; REGISTRATION NUMBER: 39,665  
; REFERENCE/DOCKET NUMBER: 600.318US4  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 612-339-0331  
; TELEFAX: 612-339-3061  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7470 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHEICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE:  
; ORIGINAL SOURCE:  
; US-08-930-285-5

Query Match 73.6%; Score 16.2; DB 4; Length 7470;

Best Local Similarity 85.7%; Pred. No. 32;

Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY --2 gtggatcacagaattgata 22

DB 3972 GTTGATGACAGAGTTGTTA 3992

## RESULT 11

US-08-695-421-5  
; Sequence 5, Application US/08695421  
; Patent No. 6268550  
; GENERAL INFORMATION:  
; APPLICANT: Gengenbach, B. G.  
; APPLICANT: Somers, D. A.  
; APPLICANT: Wyse, D. L.  
; APPLICANT: Gronwald, J. W.  
; APPLICANT: Egli, M. A.  
; APPLICANT: Lutz, S. M.  
; TITLE OF INVENTION: METHODS AND AN ACETYL COA CARBOXYLASE GENE

; TITLE OF INVENTION: FOR CONFERRING HERBICIDE TOLERANCE AND AN ALTERATION IN OIL

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.

; STREET: P.O. Box 2938

; CITY: Minneapolis

; STATE: MN

; COUNTRY: USA

; ZIP: 55402

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FASTSEQ Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/695,421

; FILING DATE: 23-AUG-1996

; CLASSIFICATION: 800

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/417089

; FILING DATE: 05-APR-1995

; APPLICATION NUMBER: 08/014326

; FILING DATE: 05-FEB-1993

; APPLICATION NUMBER: 07/917462

; FILING DATE: 21-JUL-1992

; APPLICATION NUMBER: 07/538674

; FILING DATE: 18-JUN-1990

; ATTORNEY/AGENT INFORMATION:

; NAME: Woessner, Warren D

; REGISTRATION NUMBER: 30,440

; REFERENCE/DOCKET NUMBER: 600.318US2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 612-373-6900

; TELEFAX: 612-339-3061

; TELEX:

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 7470 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; US-08-695-421-5

Query Match

Best Local Similarity 73.6%; Score 16.2; DB 4; Length 7470;

Best Local Similarity 85.7%; Pred. No. 32;

Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 gtggatcacagaattgata 22

DB 3972 GTTGATGACAGAGTTGTTA 3992

## RESULT 12

US-08-602-145-14/c  
; Sequence 14, Application US/08602145  
; Patent No. 6025336  
; GENERAL INFORMATION:  
; APPLICANT: Goltzy, Kristin L.  
; APPLICANT: Greenberger, Joel S.  
; TITLE OF INVENTION: DETERMINING EXPOSURE TO IONIZING RADIATION  
; TITLE OF INVENTION: AGENT WITH PERSISTENT BIOLOGICAL MARKERS  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk



OY 1 tggatcacacaagatttgata 22  
||| | ||||| |||||  
DB 112 TGTAGTATCACAAGGTTTGCTA 91

## RESULT 15

US-08-928-383B-1/c  
; Sequence 1, Application US/08928383B  
; Patent No. 6210921  
; GENERAL INFORMATION:  
; APPLICANT: Robert W. Pinberg, Jeffrey M. Bergelson,  
; APPLICANT: and Marshall S. Horwitz  
; TITLE OF INVENTION: CAR, A No. 6210921e1 Cocksacklevirus and Adenovirus  
; TITLE OF INVENTION: Receptor  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD, LLP  
; STREET: 28 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA: US/08/928, 383B  
; APPLICATION NUMBER: US/08/928, 383B  
; FILING DATE: 12-SEP-1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/026,100  
; FILING DATE: 13-SEP-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mandragouras, Amy E.  
; REGISTRATION NUMBER: 36,207  
; REFERENCE/DOCKET NUMBER: DFN-020  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617)227-7400  
; TELEFAX: (617)742-4214  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1584 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 60..1157  
; US-08-928-383B-1

Query Match 70.9%; Score 15.6; DB 4; Length 1584;  
Best Local Similarity 81.8%; Pred. No. 51;  
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
OY 1 tggatcacacaagatttgata 22  
||| | ||||| |||||  
DB 371 TGATGCATCCACAGATTGAGA 350

Search completed: December 15, 2001, 03:22:50  
Job time: 7565 sec

**THIS PAGE BLANK (USPTO)**







	Xhot "	
BASE COUNT	157 a	111 c
ORIGIN		101 g
		158 t

Query Match	85.58;	Score 18.8;	DB 10;	Length 527;
Best Local Similarity	90.98;	Pred. No. 1.1e+02;		
Matches 20; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0

RESULT	5
LOCUS	A1995637/c
DEFINITION	A1995637 530 bp cDNA clone GI:5842542
ACCESSION	A1995637
VERSION	A1995637.1
KEYWORDS	EST.
SOURCE	thale cress.

BASE COUNT	163 a	108 c	100 g	158 t	1 others
ORIGIN		∴			

Query Match	85.5%;	Score 18.8;	DB 10;	Length 530;
Best Local Similarity	90.9%;	Pred. No. 1.1e+02;		
Matches 20;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0

[illegible]

FEATURES	Location/Qualifiers
source	1. .608

BASE COUNT	173 a	146 c	161 g	128 t
ORIGIN				

Query Match	83.6%;	Score 18.4;	DB 11;	Length 608;
Best Local Similarity	95.0%;	Pred. No. 1.7e+02;		
Matches 19;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0

RESULT	7
CNS02BMA/c	
LOCUS	
DEFINITION	
CNS02BMA	934 bp DNA GSS 12-MAY-2000
Tetradon nigroviridis genome survey sequence PUC-Orl end of clone	
234CG22 of library G from Tetradon nigroviridis, genomic survey	

sequence.  
 ACCESSION AL190387 GI:7828491  
 VERSION AL190387.1  
 KEYWORDS GSS: genome survey sequence.  
 SOURCE Tetradon nigroviridis.  
 ORGANISM Tetradon nigroviridis  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percormorpha; Tetradontiformes; Tetradontidae; Tetradon.  
 1 (bases 1 to 934)  
 Roest-Crolius, H., Jaillon, O., Dasilva, C., Fizes, C., Fisher, C., Bouneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A. and Weissbach, J.  
 Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetradon nigroviridis  
 TITLE Unpublished  
 JOURNAL 2 (bases 1 to 934)  
 Roest-Crolius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C., Bernot, A., Fizes, C., Mincker, P., Brotlier, P., Quetier, F., Saurin, W. and Weissbach, J.  
 Human gene number estimate provided by genome wide analysis using Tetradon nigroviridis DNA sequence  
 TITLE Unpublished  
 JOURNAL 3 (bases 1 to 934)  
 Genoscope.  
 Direct Submission  
 Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases  
 This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetradon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/tetradon>.  
 FEATURES  
 source  
 1..934  
 /organism="Tetradon nigroviridis"  
 /db\_xref="taxon:99883"  
 /clone="254C22"  
 /clone\_1lb="G"  
 /note="Genoscope sequence ID : C0AC254B11SP1-end : PUC-Or1"  
 BASE COUNT 282 a 175 c 207 g 259 t 11 others  
 ORIGIN  
 Query Match 83.6%; Score 18.4; DB 13; Length 934;  
 Best Local Similarity 95.0%; Pred. No. 1.8e+02;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Oy 3 ttgtatcacagaattgatg\*22  
 Db 840 TTGCAACACAGATTTCATA 821  
 RESULT 8  
 LOCUS BF006444/c 600 bp mRNA EST 06-OCT-2000  
 DEFINITION EST434942 DSLC Medicago truncatula cDNA clone pDSL-41J15, mRNA sequence.  
 ACCESSION BF006444  
 VERSION BF006444.1 GI:10706719  
 KEYWORDS EST.  
 SOURCE barrel medic.  
 ORGANISM Medicago truncatula  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae; Medicago.  
 1 (bases 1 to 600)  
 Fedorova, M., Plesson, B. L., Samac, D. A., Gant, J. S., Vance, C. P., Gonzales, M. B. and Ellis, L.  
 ESTs from Medicago truncatula leaves and cotyledons  
 TITLE Unpublished (2000)  
 JOURNAL  
 COMMENT Contact: Deborah A. Samac

Department of Plant Pathology  
 University of Minnesota  
 495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA  
 Tel: 612 625 1243  
 Fax: 651 649 5058  
 Email: debbys@ccc.umn.edu  
 University of Minnesota name: M275482e TIGR sequence name: MTLBA567K More information is available at: <http://chrystle.tamu.edu/medicago>  
 Seq primer: SKmod (CTA gAA gTg gAT CC).  
 Location/Qualifiers  
 1..600  
 /organism="Medicago truncatula"  
 /cultivar="genotype A17"  
 /db\_xref="taxon:3880"  
 /clone="pDSL-41J15"  
 /clone\_1lb="DSL-41J15"  
 /tissue\_type="leaves and cotyledons"  
 /dev\_stage="mixture of cotyledons from five days old plants and leaves obtained from two weeks old plants"  
 /lab\_host="E. coli strain SOLR"  
 /note="Vector: pBluescript SK +/-; Site-1: EcoRI; Site-2: XhoI; cDNA was prepared from polyA+ enriched RNA from the mixture of cotyledons of five days old plants and leaves ligated into the Uni-ZAP XR vector from Stratagene and packaged using GigaPack III gold packaging extracts. Plasmids containing lambda-ZAP phage inserts were excised from the recombinant lambda-ZAP phage using Ex-Assist helper phage and propagated in SOLR cells."  
 BASE COUNT 127 a 120 c 132 g 221 t  
 ORIGIN  
 Query Match 80.9%; Score 17.8; DB 11; Length 600;  
 Best Local Similarity 90.5%; Pred. No. 3.2e+02;  
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 Oy 1 tttgtatcacagaattgat 21  
 Db 90 TGTGATCACAGATTGAGAT 70  
 RESULT 9  
 LOCUS BF006254/c 633 bp mRNA EST 06-OCT-2000  
 DEFINITION EST434752 DSLC Medicago truncatula cDNA clone pDSL-40G21, mRNA sequence.  
 ACCESSION BF006254  
 VERSION BF006254.1 GI:10706529  
 KEYWORDS EST.  
 SOURCE barrel medic.  
 ORGANISM Medicago truncatula  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae; Medicago.  
 1 (bases 1 to 633)  
 Fedorova, M., Plesson, B. L., Samac, D. A., Gant, J. S., Vance, C. P., Gonzales, M. B. and Ellis, L.  
 ESTs from Medicago truncatula leaves and cotyledons  
 TITLE Unpublished (2000)  
 JOURNAL  
 COMMENT Contact: Deborah A. Samac  
 Department of Plant Pathology  
 University of Minnesota  
 495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA  
 Tel: 612 625 1243  
 Fax: 651 649 5058  
 Email: debbys@ccc.umn.edu  
 University of Minnesota name: M275292e TIGR sequence name: MTLAU477K More information is available at: <http://chrystle.tamu.edu/medicago>  
 Seq primer: SKmod (CTA gAA gTg gAT CC).

FEATURES  
source

Location/Qualifiers  
1. .633  
/organism="Medicago truncatula"  
/cultivar="genotype A17"  
/db\_xref="taxon:3880"  
/clone="pD5LC-40621"  
/clone\_lib="DSLC"  
/tissue\_type="leaves and cotyledons"  
/dev\_stage="mixture of cotyledons from five days old plants and leaves obtained from two weeks old plants"  
/lab\_host="E. coli strain SOLR"  
/note="Vector: pBluescript SK +/-; Site\_1: EcoRI; Site\_2: XhoI; cDNA was prepared from polyA+ enriched RNA from the mixture of cotyledons of five days old plants and leaves of two weeks old plants. The cDNA was directionally ligated into the Uni-ZAP XR vector from StrataGene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-ZAP phage using Ex-Assist helper phage and propagated in SOLR cells."

BASE COUNT 148 a 132 c 124 g 229 t

ORIGIN

Query Match 80.9%; Score 17.8; DB 11; Length 633;  
Best Local Similarity 90.5%; Pred. No. 3.2e+02;  
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 tcttgatcacagaattgat 21  
|||||  
Db 188 TGTTCGATCACAATATTGAT 168

RESULT 10  
BB288727/c 286 bp mRNA EST 09-JUL-2000  
LOCUS BB288727 RIKEN full-length enriched, 2 cells egg Mus musculus cDNA  
DEFINITION BB288727 clone B020032C12 3', mRNA sequence.  
ACCESSION BB288727  
VERSION BB288727.1 GI:8989176  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 286)  
AUTHORS Kono, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Hori, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C., Kusabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata, Y., Shigemoto, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Sugihara, Y., Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomimaga, N., Toya, T., Tsunoda, Y., Watahiki, A., Watanabe, S., Yamamura, T., Yamanaka, I., Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

TITLE RIKEN Mouse ESTs (Kono, H., et al.)  
JOURNAL Unpublished (2000)  
COMMENT Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@gsc.riken.go.jp,  
URL: http://genome-gsc.riken.go.jp/  
Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S., Sasaki, N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Thermostabilization and thermoinactivation of full length trehalase and its application for the synthesis of full length

FEATURES  
source

CDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)  
Itoh, M., Katsunai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.  
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)  
Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)  
Please visit our web site (<http://genome.ritc.riken.go.jp>) for further details.

BASE COUNT  
ORIGIN

117 a 66 c 20 g 83 t

Query Match 79.1%; Score 17.4; DB 10; Length 286;  
Best Local Similarity 94.7%; Pred. No. 4.3e+02;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 gtggatcacagaattga 20  
|||||  
Db 169 GTTGCATCACAATATTGCA 151

RESULT 11  
BE581720/c 407 bp mRNA EST 09-MAY-2001  
LOCUS BE581720/k55c07.y1 TB995TM-SSR strongyloides stercoralis cDNA 5' similar to WP:119810.2 CE16413 ;, mRNA sequence.  
DEFINITION BE581720  
ACCESSION BE581720  
VERSION BE581720.1 GI:9832662  
KEYWORDS EST.  
SOURCE Strongyloides stercoralis.  
ORGANISM Strongyloides stercoralis.  
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdiltida; Panagrolaimoidea; Strongyloidea; Strongyloides.

REFERENCE 1 (bases 1 to 407)  
AUTHORS McCarter, J., Clifton, S., Chiapelli, B., Pape, D., Martin, J., Wyle, T., Dante, M., Maria, M., Hillier, L., Kucaba, T., Theising, B., Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Franklin, C., Tsagaris, V., R., Ronko, I., Kennedy, S., Maguire, L., Beck, C., Underwood, K., Steptoe, M., Allen, M., Person, B., Swaller, T., Harvey, N., Schuck, R., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.

TITLE The Washington Univ. Nematode EST Project, 1999  
JOURNAL Unpublished (1999)  
COMMENT Contact: McCarter, JP  
The Washington Univ. Nematode EST Project, 1999  
Washington University School of Medicine

## ORIGIN

	Query Match	Best Local Similarity	Matches	Conservative	Score 17.4;	DB 10;	length 646;
		94.7%			Pred. No. 4.9e+02;		
			18;	0;	Mismatches	1;	Indels 0; Gaps 0;
Qy	3	tttgatcacaagaatttgat	21				
Db	256	TTGATCACTACGATTGAT	238				

RESULT	14
LOCUS	BC448853/c
DEFINITION	BC448853 671 bp mRNA EST 16-MAR-2001 NF003C10IN1F1082 Insect herbivory Medicago truncatula cDNA clone.
ACCESSION	BC448853
VERSION	BC448853.1 GI:13367634
KEYWORDS	EST
SOURCE	barrel medic.
ORGANISM	Medicago truncatula

REFERENCE AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 671) Korth, K., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J., Flores, H.R., Imman, J.T., Weller, J.W. and May G.D.	Expressed Sequence Tags from the Sammel Roberts Noble Foundation Medicago truncatula insect herbivory library		Unpublished (2000) Contact: Korth K

University of Arkansas  
217 Plant Science Building, Fayetteville, AR 72701, USA  
Tel: 501 575 5191  
Fax: 501 575 7601  
Email: [khorth@comp.uark.edu](mailto:khorth@comp.uark.edu)  
Insert Length: 671 Std Error: 0.00  
Plate: 003 row: C column: 10  
Seq primer: TCACACAGAAAACACCTAGAC.

```

/organism="Medicago truncatula"
/db_xref="taxon:3880"
/clone="NF003C101N"
/clone_lib="Insect herbivory"
/tissue_type="local and systemic leaves"
/dev_stage="mature"
/note="Vector: Lambda Zap; Library was produced from fully
expanded M. truncatula leaves of plants fed upon by
Spodoptera exigua (beet armyworm) for 24 hours. Systemic
(un damaged leaves from injured plants) and wounded leaves
were harvested and pooled."
BASE COUNT      186 a      151 c      131 g      203 t
ORIGIN

```

Query Match	Score	DB	Length
Best Local Similarity	79.1%	17,4	671
Matches	18	Conservative	1
		Mismatches	0
		Indels	0
		Gaps	0

RESULT	15
BE283043	
LOCUS	762 bp mRNA EST 26-OCT-2000
DEFINITION	601101323P1 NCI-CGAP_Lu29 Mus musculus cDNA clone IMAGE:3493629 5'
ACCESSION	mRNA sequence.
VERSION	BE283043 BE283043.1 GI:9159023

KEYWORDS	EST.
SOURCE	house mouse.
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus 1 (bases 1 to 762)
AUTHORS	NIH-MGC <a href="http://mgc.ncl.nih.gov/">http://mgc.ncl.nih.gov/</a> .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strassberg, Ph.D.

Tissue Procurement: Gilbert Smith, Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
plate: ILAM8541 row: n column: 06  
High quality sequence stop: 661.

FEATURES	SOURCE	location/Qualifiers
		1. '762
		/organism="Mus musculus"
		/strain="CZECH II (feral)"
		/db_xref="taxon:10090"
		/clone="IMAGE:3493829"
		/clone_1lb="NCLCGAP-Lu29"
		/tissue_type="spontaneous tumor, metastatic to mammary. Stem cell origin."
		/lab_host="DH10B"
		/note="Organ: lung; Vector: pCMW-SPOK6; Site_1: Salt; Site_2: Noli; Cloned unidirectionally. Primer: Oligo dT library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"
BASE COUNT		193 a 172 c 207 g 190 t
ORIGIN		

	Query Match	Best Local Similarity	Matches	Conservative	Score 17.4;	DB 10:	length 762;
QY	1	94.7%	18	0	Mismatches	1	Indels
	1	94.7%	18	0	Mismatches	1	Indels
Db	692	94.7%	18	0	Mismatches	1	Indels
	692	94.7%	18	0	Mismatches	1	Indels

Search completed: December 15, 2001, 02:33:49  
Job time: 4989 sec

Mon Dec 17 07:48:52 2001

us-09-380-826a-7.rst

Hines  
09/380826  
Seq. IDs 1-244-704  
17Hef

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 01:19:15 : Search time 7962.77 Seconds  
(without alignments)  
3250.941 Million cell updates/sec

Title: US-09-380-826A-1  
1477  
Sequence: 1 gataatgcacgaactaac.....ccgtaaatcgatctctgacg 1477

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 17159718 seqs, 8763200856 residues

Total number of hits satisfying chosen parameters: 34319436

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_NA\_Main:\*  
1: /cgn2\_6/ptodata/2/pna/US06\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/pna/US07\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/pna/US08\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/pna/US09\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/pna/US0958\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/pna/US0959\_COMB.seq:\*  
7: /cgn2\_6/ptodata/2/pna/US0960\_COMB.seq:\*  
8: /cgn2\_6/ptodata/2/pna/US0961\_COMB.seq:\*  
9: /cgn2\_6/ptodata/2/pna/US0962\_COMB.seq:\*  
10: /cgn2\_6/ptodata/2/pna/US0963\_COMB.seq:\*  
11: /cgn2\_6/ptodata/2/pna/US0964\_COMB.seq:\*  
12: /cgn2\_6/ptodata/2/pna/US0965\_COMB.seq:\*  
13: /cgn2\_6/ptodata/2/pna/US0966\_COMB.seq:\*  
14: /cgn2\_6/ptodata/2/pna/US0967\_COMB.seq:\*  
15: /cgn2\_6/ptodata/2/pna/US0968\_COMB.seq:\*  
16: /cgn2\_6/ptodata/2/pna/US0969\_COMB.seq:\*  
17: /cgn2\_6/ptodata/2/pna/US0970\_COMB.seq:\*  
18: /cgn2\_6/ptodata/2/pna/US0971\_COMB.seq:\*  
19: /cgn2\_6/ptodata/2/pna/US0972\_COMB.seq:\*  
20: /cgn2\_6/ptodata/2/pna/US0973\_COMB.seq:\*  
21: /cgn2\_6/ptodata/2/pna/US0974\_COMB.seq:\*  
22: /cgn2\_6/ptodata/2/pna/US0975\_COMB.seq:\*  
23: /cgn2\_6/ptodata/2/pna/US0976\_COMB.seq:\*  
24: /cgn2\_6/ptodata/2/pna/US0977\_COMB.seq:\*  
25: /cgn2\_6/ptodata/2/pna/US0978\_COMB.seq:\*  
26: /cgn2\_6/ptodata/2/pna/US0979\_COMB.seq:\*  
27: /cgn2\_6/ptodata/2/pna/US0980\_COMB.seq:\*  
28: /cgn2\_6/ptodata/2/pna/US0981\_COMB.seq:\*  
29: /cgn2\_6/ptodata/2/pna/US0982\_COMB.seq:\*  
30: /cgn2\_6/ptodata/2/pna/US0983\_COMB.seq:\*  
31: /cgn2\_6/ptodata/2/pna/US0984\_COMB.seq:\*  
32: /cgn2\_6/ptodata/2/pna/US0985\_COMB.seq:\*  
33: /cgn2\_6/ptodata/2/pna/US0986\_COMB.seq:\*  
34: /cgn2\_6/ptodata/2/pna/US0987\_COMB.seq:\*  
35: /cgn2\_6/ptodata/2/pna/US0988\_COMB.seq:\*  
36: /cgn2\_6/ptodata/2/pna/US0989\_COMB.seq:\*  
37: /cgn2\_6/ptodata/2/pna/US0990\_COMB.seq:\*  
38: /cgn2\_6/ptodata/2/pna/US0991\_COMB.seq:\*  
39: /cgn2\_6/ptodata/2/pna/US0992\_COMB.seq:\*  
40: /cgn2\_6/ptodata/2/pna/US0993\_COMB.seq:\*  
41: /cgn2\_6/ptodata/2/pna/US0994\_COMB.seq:\*  
42: /cgn2\_6/ptodata/2/pna/US0995\_COMB.seq:\*  
43: /cgn2\_6/ptodata/2/pna/US0996\_COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1477	100.0	1477	17	US-09-380-826A-1
2	814.2	55.1	1536	13	US-08-965-623-1
3	814.2	55.1	1536	13	US-08-965-623-1
4	810.2	54.9	1508	15	US-09-198-953-14
5	810.2	54.9	1508	17	US-09-339-159-34
6	810.2	54.9	1508	27	US-09-694-531-14
7	808.2	54.7	1543	60	US-60-279-556-1256
8	803.6	54.4	1501	29	US-09-745-476-1
9	803.2	54.4	1501	29	US-09-745-476-1
10	803.2	54.4	1501	29	US-09-745-476-1
11	803.2	54.4	1501	30	US-09-791-592-1
12	803.2	54.4	1501	30	US-09-791-592-1
13	803.2	54.4	1501	31	US-09-821-016-5
14	803.2	54.4	1501	31	US-09-821-016-5
15	801.6	54.3	1555	9	US-08-520-946-160
16	801.6	54.3	1555	13	US-08-923-030-11
17	801.6	54.3	1555	17	US-09-349-883-11
18	801.6	54.3	1555	54	US-60-219-361-21
19	801.6	54.3	1555	54	US-60-219-361-21
20	801.6	54.3	1555	54	US-60-219-361-21
21	801.6	54.3	1555	54	US-60-219-361-21
22	801.6	54.3	1555	54	US-60-219-361-21
23	801.6	54.3	1555	54	US-60-219-361-21
24	801.6	54.3	1555	54	US-60-219-361-21
25	801.6	54.3	1555	54	US-60-219-361-21
26	801.6	54.3	1555	54	US-60-219-361-21
27	801.6	54.3	1555	54	US-60-219-361-21
28	801.6	54.3	1555	54	US-60-219-361-21
29	801.6	54.3	1555	54	US-60-219-361-21
30	801.6	54.3	1555	54	US-60-219-361-21
31	801.6	54.3	1555	54	US-60-219-361-21
32	801.6	54.3	1555	54	US-60-219-361-21
33	801.6	54.3	1555	54	US-60-219-361-21
34	801.6	54.3	1555	54	US-60-219-361-21
35	801.6	54.3	1555	54	US-60-219-361-21
36	801.6	54.3	1555	54	US-60-219-361-21
37	801.6	54.3	1555	54	US-60-219-361-21
38	801.6	54.3	1555	54	US-60-219-361-21
39	801.6	54.3	1555	54	US-60-219-361-21
40	801.6	54.3	1555	54	US-60-219-361-21
41	801.6	54.3	1555	54	US-60-219-361-21

42 789.8 53.5 1516 18 US-09-426-633A-1 Sequence 1, Appl 1  
43 789.8 53.5 1516 18 US-09-426-868-1 Sequence 1, Appl 1  
44 785.4 53.2 21083 39 US-60-061-998-618 Sequence 618, App  
45 783.6 53.1 1515 29 US-09-726-774-9 Sequence 9, Appl 1

## ALIGNMENTS

RESULT 1  
US-09-380-826a-1  
: Sequence 1, Application US/09380826A  
: GENERAL INFORMATION:  
: APPLICANT: Chappel, Rod  
: TITLE OF INVENTION: LEPTOSPIRA PATHOGENS  
: FILE REFERENCE: DATE: 79.001APC  
: CURRENT APPLICATION NUMBER: US/09/380.826A  
: PRIOR FILING DATE: 1999-11-22  
: PRIOR APPLICATION NUMBER: PCT/AU98/00145  
: PRIOR FILING DATE: 1998-03-06  
: PRIOR APPLICATION NUMBER: AU P05494/97  
: PRIOR FILING DATE: 1997-03-07  
: NUMBER OF SEQ ID NOS: 26  
: SOFTWARE: FastSeq for Windows Version 4.0  
: SEQ ID NO 1  
: LENGTH: 1477  
: TYPE: DNA  
: ORGANISM: Leptospira fainei  
: US-09-380-826a-1

Query Match 100.0%; Score 1477; DB 17; Length 1477;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1477; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gatcatgctcagaactaacgctggtcggtccttaaacatgcaagtcgagcgggtag 60  
DB 1 gatcatgctcagaactaacgctggtcggtccttaaacatgcaagtcgagcgggtag 60  
QY 61 caatccctcagcggcgaacgggtgagtaaacgttgtaactccctccgagctcggataa 120  
DB 61 caatccctcagcggcgaacgggtgagtaaacgttgtaactccctccgagctcggataa 120  
QY 121 cttccgaaaggaaagtaataccgatagtctctgttgatacaagattgataagtaa 180  
DB 121 cttccgaaaggaaagtaataccgatagtctctgttgatacaagattgataagtaa 180  
QY 181 agattattgtctgagatgagcccggtccgattagctagtgtgtgagtaattggtca 240  
DB 181 agattattgtctgagatgagcccggtccgattagctagtgtgtgagtaattggtca 240  
QY 241 ccaagcgacgactcgtgacggtcggtgagaggtgtcgcgcacaaatggaactgagac 300  
DB 241 ccaagcgacgactcgtgacggtcggtgagaggtgtcgcgcacaaatggaactgagac 300  
QY 301 ggtlccatctctacgagggagcagcagttgaagatctgtcctaattgvgggaacctgaa 360  
DB 301 ggtlccatctctacgagggagcagcagttgaagatctgtcctaattgvgggaacctgaa 360  
QY 361 gtcacactctctacgagggagcagcagttgaagatctgtcctaattgvgggaacctgaa 360  
DB 361 gtcacactctctacgagggagcagcagttgaagatctgtcctaattgvgggaacctgaa 360  
QY 361 gcaagcgacccggtgaaagaaaggttctgagattgtaagttataggcagaagaa 420  
DB 361 gcaagcgacccggtgaaagaaaggttctgagattgtaagttataggcagaagaa 420  
QY 421 ataagcagcaatgtgatactgtaactgcttaagcaccggtcctaactagctgcagcagc 480  
DB 421 ataagcagcaatgtgatactgtaactgcttaagcaccggtcctaactagctgcagcagc 480  
QY 481 cgcgtaatacgtatggtgcaagcgttctgtcgaatcatctggcgtgaaaggtgtgtag 540  
DB 481 cgcgtaatacgtatggtgcaagcgttctgtcgaatcatctggcgtgaaaggtgtgtag 540  
QY 541 cggattgtgaagcaggtgtgaaactcggtggtcaaccgtgtgctgactggaacta 600  
DB 541 cggattgtgaagcaggtgtgaaactcggtggtggtcaaccgtgtgctgactggaacta 600

DB 541 cggattgtgaagcaggtgtgaaactcggtggtcacaaccgtgtgctgactggaacta 600  
QY 601 caagtcgaggtctgagagagcgaagtgtgatactccaggtgtgacggtgaaatgtcagat 660  
DB 601 caagtcgaggtctgagagagcgaagtgtgatactccaggtgtgacggtgaaatgtcagat 660  
QY 661 atctgagagacacgaatggtgagagcgtactgtgtcctaaactgtacgtgagcag 720  
DB 661 atctgagagacacgaatggtgagagcgtactgtgtcctaaactgtacgtgagcag 720  
QY 721 aaagcgtggttagtgaacccggtattgatactcccggtatccacggtcctaactgtc 780  
DB 721 aaagcgtggttagtgaacccggtattgatactcccggtatccacggtcctaactgtc 780  
QY 781 ccaagtcgtggtgtttaaaccctcaatgaaacgaactgaagttgaagtagacgcctg 840  
DB 781 ccaagtcgtggtgtttaaaccctcaatgaaacgaactgaagttgaagtagacgcctg 840  
QY 841 ggaactatgctcgaagagtgaaactcaaggaattgacggtgtgtcgcacagcgttga 900  
DB 841 ggaactatgctcgaagagtgaaactcaaggaattgacggtgtgtcgcacagcgttga 900  
QY 901 gcaatgtgttaattcgtatgatacccaaaaacccctcgtggtgtgacatgtatcgtga 960  
DB 901 gcaatgtgttaattcgtatgatacccaaaaacccctcgtggtgtgacatgtatcgtga 960  
QY 961 tcatgtagagataatagccttcgggcagattcacaggtgtgtcatgtgtgtcgtcagc 1020  
DB 961 tcatgtagagataatagccttcgggcagattcacaggtgtgtcatgtgtgtcgtcagc 1020  
QY 1021 tctgtgtgtgagatgt 1080  
DB 1021 tctgtgtgtgagatgt 1080  
QY 1081 acctaaagttgacactgtgtgaaactcgtgtgaaactcgtgtgaaactcgtgtgaaact 1140  
DB 1081 acctaaagttgacactgtgtgaaactcgtgtgaaactcgtgtgaaactcgtgtgaaact 1140  
QY 1141 cgtcaaatccatctgagccttattgtcagagggccacacgtgtcacaatggtcgtatcag 1200  
DB 1141 cgtcaaatccatctgagccttattgtcagagggccacacgtgtcacaatggtcgtatcag 1200  
QY 1201 aggtgtcgaactcgtcgaagggagcgtatctctaaaggtgtgtgtgtgtgtgtgtgt 1260  
DB 1201 aggtgtcgaactcgtcgaagggagcgtatctctaaaggtgtgtgtgtgtgtgtgtgt 1260  
QY 1261 gttcgaactcgaactcgtcgaagggagcgtatctctaaaggtgtgtgtgtgtgtgtgt 1320  
DB 1261 gttcgaactcgaactcgtcgaagggagcgtatctctaaaggtgtgtgtgtgtgtgtgt 1320  
QY 1321 tgaatacgttccggaactgttacaacccggtcgtcacaacccgtgagttggggagcacc 1380  
DB 1321 tgaatacgttccggaactgttacaacccggtcgtcacaacccgtgagttggggagcacc 1380  
QY 1381 ggaagtggtctgttgaacgtaagagacagactaagtgtgaactcgtgaaaggggt 1440  
DB 1381 ggaagtggtctgttgaacgtaagagacagactaagtgtgaactcgtgaaaggggt 1440  
QY 1441 gaagtcgtacaaggttacgttaactcgtatcgtcag 1477  
DB 1441 gaagtcgtacaaggttacgttaactcgtatcgtcag 1477

RESULT 2  
US-08-642-229-1  
: Sequence 1, Application US/08642229  
: GENERAL INFORMATION:  
: APPLICANT: Herwig, Russell P.  
: APPLICANT: Bielefeldt, Angela R.  
: APPLICANT: Stensel, H. David  
: APPLICANT: Strand, Stuart E.  
: TITLE OF INVENTION: Degradation of Environmental Toxins by a Filamentous Bacterium



```

: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Christensen O'Connor Johnson and Kindness PLLC
: STREET: 1420 Fifth Avenue, Suite 2800
: CITY: Seattle
: STATE: Washington
: COUNTRY: USA
: ZIP: WA 98101-2347
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/642,229
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/246,865
: FILING DATE: 20-MAY-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Sheiness, Diana K.
: REGISTRATION NUMBER: 35,356
: REFERENCE/DOCKET NUMBER: UOFW19233
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (206) 682 8100
: TELEFAX: (206) 224 0779
: TELEX: 4938023
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1536 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: nucleic acid
: DESCRIPTION: "16S ribosomal DNA"
: HYPOTHEetical: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Commomonas testosteroni ATCC NO. 11996
: US-08-642-229-1

```

Query Match 55.1%; Score 814.2; DB 10; Length 1536;  
 Best Local Similarity 74.7%; Pred. No. 3.5e-251;  
 Matches 1113; Conservative 0; Mismatches 343; Indels 33; Gaps 6;

```

OY 1 gatcatggtcgaactaagcgtgagcggtgcttaacatgcaagtcgagcggtgag 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 16 GATCCTGGCTCAGATGGAACGCGCGGCGATGCTTACACATGCAAGTCGAACG 75
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 61 caa-----taactgagcggaacggtgagtgtaaacgctggttaactctcc 110
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 76 GGTCTTCGATGCTGACGAGTGCGGCAACGAGTGAATATACATCGGACGTGCT 135
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 111 tctgataactctcgaagaagaaactaataccgatagtccctgtgtatcaaga 170
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 136 TGGGGATATACACTGGAAGAGTACCTAATACCGCATGATCTACGGATGAACA 193
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 171 tgataagtaaatatgtatgtcttgagaatgagccgcggtgagctagctagt 230
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 194 GGGGACCTTCGGGCTTGTGCTACTAGACGGCTGATGCGAGATTGATGTTGG 253
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 231 taatgctacccaagcgagcgtatgacggcctggaaggtgtccggccaatgga 290
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 254 TAAAGCTTAAAGCAGCTCTCTGTACTGTGTGAGAGGAGCGACGACCACTGG 313
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 291 actgagacacgtctacatctcctaagggagcgagcttaagaattctcaatg 350
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 314 ACTGAGACACGCGCCAGACTCTCTACGGGAGGCGACAGTGGGGAATTTGG 373
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 351 aaacctgaagcagcgagcggtgagcgaagaagtgcttgcgattgtaaatg 410
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

DB 374 AAAGCTGATCCAGCAATGCCGCTGCAGATGAAGAGCCCTCGGTTGTAAC 433
OY 411 ggcaggaataaataaga-----gcaatgtatgtgtactgtccta--aag 455
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 434 GTACGGAAAGAAAGCCTGGGCTAATATCCCGGGCTATGACGGTAAAGAT 493
OY 456 caacgctactacgtgtgcagcagcccggttaatacgtatgtgcaagctgttc 515
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 494 CACGGCTAATACGTGCGCAGCAGCGCGGTAACTAGTGGGTGCAAGGCTTA 553
OY 516 taattggcgtaaaagggtgctgtagcggaattgtgaatcaggtgtgaacac 575
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 554 TTACTGGCGGTAAAGCGGTCCGACGCGGTTTGTAGACAGTGGTAAATCC 613
OY 576 aaccgtgtgcctgcacttgaactaagaatctcggaattgggagagcgagga 635
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 614 AACCTGGGAACCTGCCATTTGTACTGCAAGGTGAAGATGGCGAGAGGGAT 673
OY 636 aggtgagcgtggaatgtgtagatactcggaggaacaccagttggcgagcg 695
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 674 GCGTGTAGCAGTGAATGCGTATGTCGGAGGAACACCGATGGGAGGCAAT 733
OY 696 ggtcaaaacttgacgtctgagcagaagcgttggtagtaaacggatagata 755
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 734 GGGCGTGCATGACGCTCATGCACGAAGCGTGGGAGCAACAGGATTAAG 793
OY 756 taatcagcgccctaaacgttgtctaccagttgttgggggttttaaccct 815
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 794 TATGTCACCCCTTAAACGATGTCACACTGGTGTGGGTCTTAACATGAC 853
OY 816 cttaaggttaagtagaacgccttgggggagctatgctcgaagaagtgaac 875
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 854 CTACCGCTGAAGTTACCGCCTGGGAGTAGCGCCGCAAGGTGAATCAAG 913
OY 876 gacgggggttcgcgacaagcgttgaagcatgtgttaactcgtatgatacc 935
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 914 GACGGGAGCCCGCACAAAGCGGTGATGATGGTTAATTCATGCAACGCA 973
OY 936 caacttggcttgacatgatactgaatcagatgagatataatgaccttgg 991
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 974 TACCCACTTTGACATGGAGAGAACTTACAGAGATGTTGGTGTGCAAA 1033
OY 992 -ttcaaggtgtcgtcagttgtcgtcaagctcgttgcgtgagatgttgg 1050
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1034 GCACACAGGTGGCTGATGGCTGTGCTGACGTGTGTCGAGATGTGGTTAA 1093
OY 1051 caagcagcgaacccctatcgtatgtcttaccttaagtgttgacgtgtg 1110
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1094 CAACGAGCCCAACCTTGGCATTAGTTGCTACATTCAGTTGAGCACCT 1153
OY 1111 cgttgaacaaacggaggaagcggtgaatgaatcctcaatgtgacctt 1170
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1154 CGGTGACAAACCGGAGGAAGTGGGATACGTCAAGTCTCTATAGTGG 1213
OY 1171 gccacacacgtgtcacaatggtccgatacagaggtgtcgaactcgaaga 1230
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1214 GCTACACAGCTCATTAATAGCTGCTGATCAAAAGGTTGCCAACCCGGAG 1273
OY 1231 ctctaaagtcggtccagttcgatgttggtgtcgaactcgaacccatgaa 1290
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1274 CCGATTAAACCGAGTCTTAATCCGATCGCATTCGCACTCGATCGAAGT 1333
OY 1291 cgttagtaatcggagtagaacgacgcggtgaatcgttccggaccttga 1350
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1334 CGCTAATATCGGTGATCAGAAATGTACGCTGATACGTTCCGGGTCTT 1393
OY 1351 cccgtcaccaacactgtaggggagcaccgagtggtcttctgtacacgta 1410
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1394 CCGGTACACACATGAGGAGGGGTCTGCCAGAAAGTAG-6TAGCTTAACG 1452
OY 1411 gactactaaggtgaaactcgtlaaagggttgaagtcgtatacaagtaac 1459
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1453 GCTTACCACGCGGGGCTGCTGACTGGGGTGAAGTGTAAACAAGGTAGC 1501

```

```

RESULT 3
US-08-965-623-1
: Sequence 1, Application US/08965623
: GENERAL INFORMATION:
: APPLICANT: Bielerfeldt, Angela R.
: APPLICANT: Stensel, H. David
: APPLICANT: Strand, Stuart E.
: APPLICANT: Herwig, Russell P.
: TITLE OF INVENTION: Degradation of Environmental Toxins by a
: TITLE OF INVENTION: Filamentous Bacterium
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Christensen O'Connor Johnson and Kindness PLLC
: STREET: 1420 Fifth Avenue, Suite 2800
: CITY: Seattle
: STATE: Washington
: COUNTRY: USA
: ZIP: WA 98101-2347
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/965,623
: FILING DATE: November 6, 1997
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/246,865
: FILING DATE: 20-MAY-1994
: APPLICATION NUMBER: US 08/599,867
: FILING DATE: 12-FEBRUARY-1996
: APPLICATION NUMBER: US 08/642,229
: FILING DATE: 30-APRIL-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Sheiness, Diana K.
: REGISTRATION NUMBER: 35,356
: REFERENCE/DOCKET NUMBER: UOFW11488
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (206) 682 8100
: TELEFAX: (206) 224 0779
: TELEX: 4938023
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1536 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: nucleic acid
: DESCRIPTION: "16S ribosomal DNA"
: HYPOTHEtical: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Comamonas testosteroni AFCC No. 11996
: US-08-965-623-1

```

```

Query Match 55.1%; Score 814.2; DB 13; Length 1536;
Best Local Similarity 74.7%; Pred. No. 3, Se-251;
Matches 1113; Conservative 0; Mismatches 343; Indels 33; Gaps 6;

```

```

QY 1 gatcattgctcagaactaagcgtgagcgctgcttaaacatgcaagtcgagcggtgtag 60
DB 16 gatcctgctcagatgacgctgagcgctgcttaaacatgcaagtcgagcggtgtag 75
QY 61 caa-----taactagcgcgagcgaggtgtagtaacacgtgtgtaactctcccgag 110
DB 76 gctgctgctcagatgacgctgagcgagcgagcggtgtagtaacacgtgtgtagtaac 135
QY 111 tctgagtaactctcgaagaaggaagcctaatacgcggtgtagcccggtgtagtaacagatt 170
DB 111 tctgagtaactctcgaagaaggaagcctaatacgcggtgtagcccggtgtagtaacagatt 170

```

```

DB 136 tgggggataactactcgaaagagtagtaataaccgcatgacgatacgaataagca-- 193
QY 171 tgatagtgtaagattatctgttgagatgagcccggtgtagttagtggtag 230
DB 194 ggggaccttgcggcccttgctactagagcggtgagtgagatgagttaggtggtggg 253
QY 231 taatggtcaccagagcgagcgtgtagccgctggaaggtgctgcgcgaacaatgga 290
DB 254 taaagccttaccagacgctgtagctgtagctgtagagagacgacgacgacgacg 313
QY 291 actgagacaggtccatactctcagcggagcagcagtgtaagaaactgtcctaattgg 350
DB 314 actgagacaggtccatactctcagcggagcagcagtgtaagaaactgtcctaattgg 373
QY 351 aaacccctggaagcagcagcgtgtagcgaagaaggtcctgtagtaagtccta 410
DB 374 aaacccctggaagcagcagcgtgtagcgaagaaggtcctgtagtaagtccta 433
QY 411 ggcaggaataaagca-----gcaatgtagttagtactgctccta--aag 455
DB 434 gtacggaacgaataagcctgggctaatatcccggtcagcgtacggtacggaataag 493
QY 456 caccggtcactacgtgtagcagcagcgtgtagcagcagcgtgtagcagcagcgtg 515
DB 494 caccggtcactacgtgtagcagcagcgtgtagcagcagcgtgtagcagcagcgtg 553
QY 516 tcaatggtggtgtagcagcagcgtgtagcagcagcgtgtagcagcagcgtgtagc 575
DB 554 ttactggtggtgtagcagcagcgtgtagcagcagcgtgtagcagcagcgtgtagc 613
QY 576 aaacccgtggtgtagcagcagcgtgtagcagcagcgtgtagcagcagcgtgtagc 635
DB 614 aaacccgtggtgtagcagcagcgtgtagcagcagcgtgtagcagcagcgtgtagc 673
QY 636 aggtggtggtgtagcagcagcgtgtagcagcagcgtgtagcagcagcgtgtagc 695
DB 674 gctggtggtgtagcagcagcgtgtagcagcagcgtgtagcagcagcgtgtagc 733
QY 696 ggtcctaaacgtgtagcagcagcgtgtagcagcagcgtgtagcagcagcgtgtagc 755
DB 734 ggtcctaaacgtgtagcagcagcgtgtagcagcagcgtgtagcagcagcgtgtagc 793
QY 756 taatccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 815
DB 794 taatccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 853
QY 816 ctacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 875
DB 854 ctacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 913
QY 876 gacggtggtgtagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 935
DB 914 gacggtggtgtagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 973
QY 936 caactggtgtagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 991
DB 974 taacccacttggatgacgagcagcagcagcagcagcagcagcagcagcagcagc 1033
QY 992 -ttcagaagtgtagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1050
DB 1034 gcacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1093
QY 1051 caacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1110
DB 1094 caacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1153
QY 1111 cgtgtagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1170
DB 1154 cgtgtagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1213
QY 1171 gccacacgtgtagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1230
DB 1214 gctacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1273

```

QY 1231 ctctcaaaagctcggtlcccaagttccgatttggggtcttgcgaactcgacccttgaaagtcggaat 1230

Db 1274 CCCCTAAAGGCCAATCTGTAATGCGGAAATCGGAAGTCTCGAACTTCACAGCCGTGAAGTCGGAAAT 1333

QY 1291 cgtcagtaatcgcggatcgaacagctgcgcgggttgtaatacgtctccgcgaactctgtacacacg 1350

Db 1334 CGCTAGTAATCTGTGGATCGAATAATGTACACGGTGAATACGTTCCCGGGCTTGTATACACACCG 1393

QY 1351 cccgctcacaccacttgatggggagaccccgaaatggtctctgttcaaccgtaaggagaca 1410

Db 1394 CCCGTCACACCAATGAGGAGCGGGCTCTGCACAGAAGTAG -GTACCTTAACCGTAAGAGGAGCG 1452

QY 1411 gactcactcaagtgtaaacctcgttaagggggttgaaagtcgttaacaaggtacc 1459

Db 1453 GCTTATCACACGCGGGGTTCTGTGACTGTGGGTGAAGTCGTATCAAGAAGTACC 1501

```

US-09-198-955-14
US-RESULT 4
Sequence 14, Application US/09198955
GENERAL INFORMATION:
APPLICANT: Andersen, Lene Nonboe
APPLICANT: Schulteln, Martin
APPLICANT: Lange, Niels Erik Krebs
APPLICANT: Bjornsvad, Mads Eskelund
APPLICANT: Moeller, Soeren
APPLICANT: Glad, Sanne O. Schroeder
TITLE OF INVENTION: Novel Peptate Lyases
FILE REFERENCE: 5378.200-US
CURRENT APPLICATION NUMBER: US/09/198,955
CURRENT FILING DATE: 1998-11-24
EARLIER APPLICATION NUMBER: 1343/97
EARLIER FILING DATE: 1997-11-24
EARLIER APPLICATION NUMBER: 1344/97
EARLIER FILING DATE: 1997-11-24
EARLIER APPLICATION NUMBER: 60/067,249
EARLIER FILING DATE: 1997-12-02
EARLIER APPLICATION NUMBER: 60/067,240
EARLIER FILING DATE: 1997-12-02
EARLIER APPLICATION NUMBER: 09/073,684
EARLIER FILING DATE: 1998-05-06
EARLIER APPLICATION NUMBER: 09/184,217
EARLIER FILING DATE: 1998-11-02
SOFTWARE: FastSeq for Windows Version 3.0.0
SEQ ID NO 14
LENGTH: 1508
TYPE: RNA
ORGANISM: Bacillus sp.
US-09-198-955-14

```

Query Match	54.9%	Score 810.2	DB 15	Length 1508
Best Local Similarity	61.7%	Pred. NO. 6.9e-250		
Matches 911	Conservative 220	Mismatches 303	Indels 49	Gaps 10

[illegible]

QY	232	aatggtcacaagaagcgacatctgtaacgagctctgaagaggctgcacggacaataagg	291
Dd	241	aacggcunuaacaagcgacagaaugcguaagccgaacuuagaggggugacgcacacuggga	300
QY	292	ctgagacacggtctcaatactctcaacggagagcgacagtctgaagatcttgctcaatgvggga	351
Dd	301	cugagacacggcccgacgacuuacgaggagcgacguagagggaatuuucgcgaauagacga	360
QY	352	aaccctgaagcagcgacggccgctgaaacgaagaaggtctctgatttgtaaaagtctcatag	411
Dd	361	aagucuaacgaggaagcgaacggcgcgugagugagaaaggguuucgucuguaaagcunugug	420
QY	412	gcaggaagaaataagacagcaatctgtaatgtaactctgcgtc-----aaag	455
Dd	421	uuaagggaagaacaaugugcguaaataaunaggcggcaccuagacguuacuuacaagaag	480
QY	456	caccggcctaaactacgtctgcacagcgacggtaataactgatagtgtgcaagcgttctcgaa	515
Dd	481	ccaacgcuuaacaatagcugtcacagcgccgcuuaaactgaaugugcgaaguguuacggaa	540
QY	516	tcatctgggcgttaaaagggtctgctgagcgagatttgtaagtaagggtgtgaacaacggcgga	575
Dd	541	uuauuggcgtuaaagcgcgcgacgagcguguuuuuaagucuaagugaaacuuacggcguc	600
QY	576	aaccggctggcccttgacactctgaacatacacaagctctggaattctggagagagcgaaagtgaatcc	635
Dd	601	aaccgcgagcguguaaagucuaacugcgagacuuagaaacagaagaagagagugaaauuc	660
QY	636	aggtgtgtagcgttgaatgatcgtctagataatctgtagaagaaacacgaatctgacgaagcgtacttgc	695
Dd	661	acguuguaagctggugaaagucagaaatuauguaagaaacacccagugcgaaaggcgacacucuu	720
QY	696	ggctcaaaactgaacgctcgaaagcagcaagagctgggtcagttaaacggagatctgataccggg	755
Dd	721	gguucuguaacugacugcugagcgcgacgaagaagcgugggagacaaacaggaauaagaaucug	780
QY	756	taattccacgcccataacgctgtgtctcaactgtgttgagggtcttaa--ccctcagttaacga	813
Dd	781	uauguccaagcgguaaacaagaaugaguguaaguguaaggguuauucgaaucgucuaugucgga	840
QY	814	aactcaacgatactaaagttagacagccgcttggagactatgtctgcaagagatgaaacccaagaa	873
Dd	841	aguuacaacaauaagacacucgcgcgcuguguaagcgcgcgaadguuugaaacuacaagaagaa	900
QY	874	ctgacgggggtctcgacacaaagccggctgtagacatgtgtgtcttaattctgaatgataccccaagaac	933
Dd	901	uugagagggggccgcgcacaaagcagugagcgacuguguuuaucgaaagcgaacggaagac	960
QY	934	ctcacccctgggctctgaatctgatatctgaatactgtagagata--tatgacgtctcgggcaga	991
Dd	961	cuuacacgaugucuaagacauccuuagacacucuaagagaaugaaugaauguuuccuucggggagaca	1020
QY	992	t---tcaacagctgtctgatactgtctgtctgcagccgctgtcgtctgagatactgtggcttlaagtc	1048
Dd	1021	uaagugaaacaaagugugcagugaugugucugacgucgcugugugagaaugaauguuuguaaguc	1080
QY	1049	cgcacaacggcgcaaacccct-atcgtatgtgtctacttaagtgtggcactgtgtaacgaaac	1107
Dd	1081	cgcacaacggcgcaaaccccuuauguaucuaagucgcagcauuuugugggcacaucuaaaguguaac	1140
QY	1108	tgcacgttgaacaacccgagagaaagcgagggaatgacttcaaatctctatgaccttatactgttgc	1167
Dd	1141	ugccgguguaaataacccggaaggaagguuggggaaacgynucaaaucuaucuaugcccuuaugacc	1200
QY	1168	aggggcacacacgttgcatacaatgtggcgatacaagaaggtctgcacaactcgcgaagaagagact	1227
Dd	1201	uuggcuaacacacgugucuaacaagaauguguaacaagaagcagcaaaacgcggagugacgacc	1260
QY	1228	aactcctaaaggtctggctccgaactctcgatattgggtgtctgcaactcgaaaccccatgaaagtct	1287
Dd	1261	aaucucuaaagccaauucacgaaugucggaauuguaugcgucacacucgcuaacuaagaagccgg	1320
QY	1288	aatcgctagtatactcgatatacgtatccgcggctgataactgttcccgggaccttctgacaca	1347









```

: TITLE OF INVENTION: Methods for Monitoring Multiple Genes
:
: TITLE OF INVENTION: Expression
:
: FILE REFERENCE: 10085.000-US
:
: CURRENT APPLICATION NUMBER: US/09/880,598A
:
: CURRENT FILING DATE: 2000-10-06
:
: NUMBER OF SEQ ID NOS: 1662
:
: SOFTWARE: FastSeq for Windows Version 4.0
:
: SEQ ID NO 1291
:
: LENGTH: 2484
:
: TYPE: DNA
:
: ORGANISM: Bacillus licheniformis
:
: FEATURE:
:
: NAME/KEY: misc_feature
:
: LOCATION: (1)...(2484)
:
: OTHER INFORMATION: n = A,T,C or G
:
: US-09-680-598A-1291

```

Query Match	54.4%	Score 803.6;	DB 27;	Length 2484;
Best Local Similarity	75.4%;	Pred. No. 1.2e-247;		
Matches 1134;	Conservative	0;	Mismatches 324;	Indels 46;
				Gaps 9

OY	1	gatactggcccaagaactaaagccgagcgagcgtcttaaacatggaagctcgaacgaggtag	60
Db	2180	gattcttggtctcagagcagaaacgctggccgctgacctatracatgcaagctgaacgaccca	2121
OY	61	caatlaact-----agcgagcaagcggtgtaacacgt..gtaatcttcc	105
Db	2120	cggagacttgctcccttaggtcagcgcgagcgaggtgaatacactggtgaacttggct	2061
OY	106	cgaagctcgtgataaacttcgcgaagaagataaccggaagt-----ctgct	157
Db	2060	gtaagacttggaataactccgggaaacggggcgttaataccggaatgcttgattgaacccgact	2001
OY	158	ggaatcaagaatttgtagtaagaatttatctgcttggaatgaaagcccgccgagatt	217
Db	2000	gctttcaatctataaagctgcttttagcttaaccattacagatgaccccgccgacttag	1941
OY	218	ctagcttgtagaagtaatggtctcaagaagcgagatcgtatgcgagctgaaaggtgctc	277
Db	1940	ctaggttgtagaagtaacggcttccaaaggcgaacgagatcgcttaccacactgaagggat	1881
OY	278	cggccacaatggaacttgaagacabggtcccaactcctaaggagcgagcaagttaaagatct	337
Db	1880	cggccacactcgggaactgagacagcggcccgacactctacggagcagcagtagaggaact	1821
OY	338	tgctcaatctgggggaaaaacctgaaagcgagacgcgcgtgaaagaaaggtcttcgact	397
Db	1820	tcccaatggagacaaagctctgacggagcaacggcggtgaagtaagattttggagct	1761
OY	398	gtaaggtcattcggcagggaaaaataagacgaagtgtgatagttaccgtctgctct-----	451
Db	1760	gtaaaactctggtgttaagggaaabaacaagtaaccttgcgaatagggcggactttacgct	1701
OY	452	-----aaagcacggtactaactcgtgcacgacgcgcggtgaatacgtatgtgca	501
Db	1700	acctaaaccaaagaaagcagcagcgtactaaactgacgacggcggaatgactgaagtgca	1641
OY	502	agcgttgctcgatcatcttggcggtaaaggtggtgtagcggaatttgttaagcaggtg	561
Db	1640	agcctttctcggaaattattttgggggttaaacgcggccgaggggtttcttaagcttgaatg	1581
OY	562	aaaactcgggctcaaacccgtgacctgcaacttgaactcaagctcgaagttctggagag	621
Db	1580	aaagcccccgggctcaacccgggaggggtcatttggaacctggggaacttagtgcacaaag	1521
OY	622	gcaagtgaatctccaggtgtagcgtgtgaatctgctgatatcttggagaacacaggtgc	681
Db	1520	gagagtggaatttcacggtgaacgggtgaatgctgagatgttgaggaacccagtgctc	1461
OY	682	gaagcgactctgctcacaactgaagcttgagggcagaagaagcgttggttagtaacggt	741
Db	1460	gaagcgactctctggtctgttaactacgcttgagggcgaagagcgtgggagaccgaagag	1401

QY	742	atlagataccocggtaa	tcacgcgcctaaacgtgtgcataccaggttglttgggggtttt--a	799
Db	1400	ATTGATTAACCCCTGGTACTCCAGCCCGTAACCGATGAGTGTCTAAGTATTAGAGGGTTTCCG	1341H	
QY	800	accctcagtaa	gaacctaaacggatlaagtataccgcctcggggagctatgcttcgcacaagat	859
Db	1340	CCCTTATGTCGTGTGAGCAAAACGATTAAAGACTCCGCTGGGGAGTAAAGTTCGAAAGCT	1281	
QY	860	gaacctcaaa	ggaatltgaacgggggtccgcacaacgggtggagcatgtggtttaa	919
Db	1280	GAAACTCAAAAGAAATTATGACGGGGGCCCGCAACAAGCCGTGAGACATGTGGTTTAATTCGAA	1221H	
QY	920	gatacccccaaaa	acctactactgtggttgatc-ggatcgaatactgtlaagatatata	978
Db	1220	GCAACGGGAAGAACCTTATACCAAGGTCTTTGAAATCTCTGCAACCCCTAGAGATAGGGCTTC	1161	
QY	979	gacctc--gggcagat	ctacacagtgctgcatagtgtgcataacgcgtgtcgtgaatgt	1036H
Db	1160	CCCTTCGGGGGGCAAGTGAACAGGTGTGATGATGTTGTCTACGTCTGTCTGTAGATGT	1101	
QY	1037	ttgggtlaagt	ctccgcacaacgacgaacccct-atcglatgtgtcactaa	1095H
Db	1100	TGGGTATAGTCCCGCAACGAGCCCAACCCCTTGATCTTATGTGCCAGCATTAAGTTGGGCA	1041H	
QY	1096	ctgtgaagaac	atgcgtgtacacacacgcggagagacggggttgatgtaactctc	1155H
Db	1040	CTTAAAGTATGATCTCCGCTACCAAAACCGGAGAGAGGTGGGATGACGTCAAAATCATATG	981	
QY	1156	gacctatgtc	acagggcacacacagctgtcactaaatggccgatacaagggctgcacaacg	1215H
Db	980	CCCTTATGATGACTGGGCTAACACAGTGTCAATATGGGCAAGAACAAAGGCCAACGAAGCG	921	
QY	1216	caagaggagga	ctaaatctctaaagtcggtcccagttcggaatgtgggtctgcacatcgacc	1275H
Db	920	CGAAGCTAAGCCCAATCCCACAAATCTGTTCTCAGTTCCGATGCGCATGTGCAACTCGACT	861	
QY	1276	ccatgaagtc	cgatcgtctatctcgcgagatcagatatgcgcggttgataatgctccgg	1335H
Db	860	GCGGAGAGCTGGAAATCCGTAGTAATCGCGGATACGATGATCCCGCGGTAAATACGTTCCGG	801	
QY	1336	acctgtgata	caacccgcgtcacacaacctgagttggggagcaaccgaaatgtgcttctgt	1395H
Db	800	GCCTTGTATCACACCGCCCGTACACACAGAGATTTGTATACACCCGAACTCG-GTAGAGT	742	
QY	1396	aacctgaagg	agacagactactaagatltgaacatcgtcaaaaggggttgaagtcgtlaacaag	1455H
Db	741	AACCTTTTGGAGCCAGCCCGCGCAAGGTGGGACAGATGATTTGGGGTGAACCTGTAACAGG	682	
QY	1456	tacc	1459	
Db	681	TAGC	678	

```

RESULT      9
US-09-745-476-1
; Sequence 1, Application US/09745476
; GENERAL INFORMATION:
; APPLICANT: CANON INC.
; TITLE OF INVENTION: Preparation of Poly-hydroxyalkanoic Acid
; FILE REFERENCE: 4351008
; CURRENT APPLICATION NUMBER: US/09/745,476
; CURRENT FILING DATE: 2000-12-26
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: Microsoft Word
; SEQ ID NO 1
; LENGTH: 1501
; TYPE: DNA
; ORGANISM: Pseudomonas jessenii P161 ; FERM P-17445
US-09-745-476-1

```

Query match 54.4%; Score 803.2; DB 29; Length 1501;

Best Local Similarity 75.9%; Pred. No. 1.2e-247;  
Matches 1121; Conservative 0; Mismatches 318; Indels 37; Gaps 9;  
QY 18 aagcgtgagcgccgctttaaactgaaagtcgagcgg-----gtagcata 65  
Db 3 aagcgtgagcgccgctttaaactgaaagtcgagcggatacgagcgagcgttgcctcgaa 62  
QY 66 cctagcggcgaacgggagaaacacgtagtaactctccctccgagctcgtggatacttc 125  
Db 63 ttcagcggcgagcgaggtagtaactgtagtaactcgtgtagtgagggaacaagctc 122  
QY 126 cgaaggaagacataaccgagtagctcgttgcatacaagatttgataggttaagatt 185  
Db 123 cgaaggaagacataaccgagtagtaactcgttgcatacaaggaagggagacttcggcc 180  
QY 186 tattgcttgagatgagcccgccgagcttaagctagcttgtagtaagtaagctacaag 245  
Db 181 ttcgctatacagatgagccctaggtcgtatgactagcttgtagtaagtaagctacaag 240  
QY 246 gcgacgatacgtacggccgtgagaggtgtccggccaatggaaactgagacagctcc 305  
Db 241 gcgacgatacgtacggccgtgagaggtgtccgacactggaactgagacagctcc 300  
QY 306 atactctacggagggagcagcaggttaagaatctgtcatactgggggaaacctgaagcagc 365  
Db 301 agactctacggagggagcagcaggttggaataatgacaatggggcgaagcctgataccagc 360  
QY 366 gacgcgcgtggaacgaagaggtcttcgagattgtaagttcattcaggcgaagaaataag 425  
Db 361 cagcgccggtgtgtggaagaggtcttcgagattgtaagttcattcaggcgaagaggg 420  
QY 426 cagcaaa-----tgtagatgtgtaactgctca--aagcaacggagctaacag 470  
Db 421 catlaaacctlaacgtagttagttagttagttagttagttagttagttagttagttagttag 480  
QY 471 tgcgacgagcgcgcgttaatacgtatgtgtgcaagcgtgtgttcggaatactatggggagtaag 530  
Db 481 tgcgacgagcgcgcgttaatacgtatgtgtgcaagcgtgtgttcggaatactatggggagtaag 540  
QY 531 ggttgctgtagcgagattgttaagtcagtgtagtaaaactgagcggtcacaacccgtgctgca 590  
Db 541 cgcgcgtgaggtgtgtgttagttagttagttagttagttagttagttagttagttagttag 600  
QY 591 ctggaactcaagctcgtgaggttgggagggagcgaagtggaattccaggtgtgaaggtgaa 650  
Db 601 tcaaaaactcaacagctcagagtagttagttagttagttagttagttagttagttagttagt 660  
QY 651 atcgtagatatactggaaggaacacacagtgagcgaagcgaactgtgtgctcaaaactagc 710  
Db 661 atcgtagatatactggaaggaacacacagtgagcgaagcgaactgtgtgctcaaaactagc 720  
QY 711 ctgagagcagaaagcgtgtgtagtaaaacgggattagatacccggtatccacgcccctaa 770  
Db 721 ctgagagcagaaagcgtgtgtagtaaaacgggattagatacccggtatccacgcccctaa 780  
QY 771 acgttgctacacagctgtgtgg--ggtttaacctcaagtaacgaactaacgaattagc 829  
Db 781 acgttgctacacagctgtgtgg--ggtttaacctcaagtaacgaactaacgaattagc 840  
QY 830 agaacgccttgaggagctatgctcgcgaagtgaaactcaaggaattagcggggttcgca 889  
Db 841 tgaacgccttgaggagctatgctcgcgaagtgaaactcaaggaattagcggggttcgca 900  
QY 890 caagcgttgagagcgtgtgtttaattcgaatgatacccaaaaactcaacgtgggcttgac 949  
Db 901 caagcgttgagagcgtgtgtttaattcgaatgatacccaaaaactcaacgtgggcttgac 960  
QY 950 at--ggaactcaatcatgtagagataatgagccttcgggcagatt--cagaagtgtcgca 1006  
Db 961 atccaagtaacttcagagataatgtagtgccttcggaacttgagacaggtgtcgca 1020  
QY 1007 tggttgctgcagctgctgtctgtagagattgtgggtlaagctccgcaacgaagcaacccc 1066  
Db 1007 tggttgctgcagctgctgtctgtagagattgtgggtlaagctccgcaacgaagcaacccc 1066

Db 1021 tgcgtgctgaagctcgtgtctgtagagattgtgggtlaagctccgtaacgaagcgaacct 1080  
QY 1067 tatcgtagttagc---taccttaagttgggacactgtagcaaaactgcccgtggaacaaacg 1123  
Db 1081 tgccttaagtttaacagcagcgttaatgtgtggacacttaaggaagactgcggtggaacaaacg 1140  
QY 1124 gaaggaagcgaggagtagcgtlcaaaatccctcaatgagccttaatgltccagggccaacagctgc 1183  
Db 1141 gagggaaggtggggatgagcgtcaagttcaatgagccttaagccttcaggccttggtctacacagctgc 1200  
QY 1184 taacatgacgagatacaagaggttcgccaactcgcaaggggagagccttaactctcaaaagctgcg 1243  
Db 1201 taacatgacgagatacaagaggttcgccaagcgaagctgcaagcttaactcccaaaacgca 1260  
QY 1244 tccagatcgagattggggtctgcgaactcgcaacccatgaagtcgaatcgtagtaacgc 1303  
Db 1261 tgcgtagtcggagatcgacgtctgcgaactcgagctggtgaagtcggaatcgctagtaacgcg 1320  
QY 1304 ggalcaagcagtcgagctggaatacgttcccgagccttgtaacacacgcccgtlcaacacac 1363  
Db 1321 gaalcaagatgtcggtgtaatacgttcccgagccttgtaacacacgcccgtlcaacacac 1380  
QY 1364 ctgagtgaggagaccccggaagtggtcttgttaacggtlaaggaacagacttaaggtg 1423  
Db 1381 gggagtggtgtgaccccggaagtag--ctagttcaactctcgaggagcgttatcacaggtg 1439  
QY 1424 aaactcgtaaggggtgaaagtcgtaaacaggtacc 1459  
Db 1440 tgatctagctggtgggtggaagtcgtacaaaggtagc 1475

RESULT 10  
US-09-748-205-1  
; Sequence 1, Application US/09748205  
; GENERAL INFORMATION:  
; APPLICANT: Canon, Inc.  
; TITLE OF INVENTION: Polyhydroxyalkanoate its manufacturing method, and microorgan  
; TITLE OF INVENTION: Those are used for the method.  
; FILE REFERENCE: 4351009  
; CURRENT APPLICATION NUMBER: US/09/748, 205  
; CURRENT FILING DATE: 2000-12-27  
; NUMBER OF SEQ ID NOS: 1  
; SEQ ID NO 1  
; LENGTH: 1501  
; TYPE: DNA  
; ORGANISM: Pseudomonas jessenii 161 strain.  
US-09-748-205-1

Query Match 54.4%; Score 803.2; DB 29; Length 1501;  
Best Local Similarity 75.9%; Pred. No. 1.2e-247;  
Matches 1121; Conservative 0; Mismatches 318; Indels 37; Gaps 9;  
QY 18 aagcgtgagcgccgctttaaactgaaagtcgagcgg-----gtagcata 65  
Db 3 aagcgtgagcgccgctttaaactgaaagtcgagcggatacgagcgagcgttgcctcgaa 62  
QY 66 cctagcggcgaacgggagaaacacgtagtaactctccctccgagctcgtggataacttc 125  
Db 63 ttcagcggcgagcgaggtagtaactgtagtaactcgtgtagtgagggaacaagctc 122  
QY 126 cgaaggaagacataaccgagtagctcgttgcatacaagatttgataggttaagatt 185  
Db 123 cgaaggaagacataaccgagtagtaactcgttgcatacaaggaagggagacttcggcc 180  
QY 186 tattgcttgagatgagcccgccgagcttagctagctgtgtaggttagttagttagttagttag 245  
Db 181 ttcgctatacagatgagccctaggtcgtatgactagcttgtaggttagttagttagttagttag 240  
QY 246 gcgacgatacgtacggccgtgagaggtgtccggccaatggaaactgagacagctcc 305  
Db 241 gcgacgatacgtacggccgtgagaggtgtccgacactggaactgagacagctcc 300



OY	306	atactcttaaggagagcagcagatcgaagaaattctgtctcaaatgggaggaaaccttgaaagcgc	365
Db	301	agactcttaaggagagcagcagatcgggaaatatttttttttttttttttttttttttttttt	360
OY	366	gaacccgcgtgaacaggaagaggtctccgatatctgaagtctactgagcaagaaataaag	425
Db	361	catgcgcgctgtgtgaagaagagctcttcgatatctgtaagaaccttgaattctggaggaaagg	420
OY	426	caagaca-----tgtgatgatgtactctgtccta--aaagcccgctactacgc	470
Db	421	cattaaactataactgtatagtgcttttgaagctttaaocgcagataaagcacccgctaaactcg	480
OY	471	tgcacagacccgcggttaatacgttatatgtgcagaacgctgtctcggaaatctatggcgctaaag	530
Db	481	tgcacagacagccgcgttaatacagaaggtgtccaaacgcttaactcgaaattactgtgcgttaaag	540
OY	531	gtgtcgtcagaacgcatctgttaagtcacagctgtgtaaaactgggggtctaaacccgttgccgtca	590
Db	541	cgcgcgtcagaatgtgttctgttaagcttctgaatgtgtgaagcccggtctcaacccgtggaaactgc	600
OY	591	cttgaataactcaaaagtctcgtgaagcttctgggaagagcgaaatggaaattccaaagtgtgaacgttga	650
Db	601	ttcaaaactgcaagcctaagaggtatgttagaggggtgggtggaattctcctgtgttagcgtgtga	660
OY	651	atgtgtatgatatctcgtgagagaaacacccgctgtgcgaagcgacatgcctgcgtctcaaaactgcg	710
Db	661	atgtgtatgatatagaaagagaaacacccgctgtgcgaagcgacacccctgcgtctgaatactgaca	720
OY	711	cttgagcagcaaaagcgtgtgggttagttaaocgggtatgaatacccggttaataccacgctctaa	770
Db	721	cttgaggtgtcgaagaagcgtgtgggtgaagcaaacggtatgaataccctgtgttagtccacgctgtaa	780
. OY	771	acgttcttactcaagtttctgtgg--ggttttaaacctctaagtaacgaacctcaacgaaatgaat	829
Db	781	acgattgtcaactaactacgcttctggagcccttgagctctttagtggcgacgactaaacgactaaagt	840
OY	830	agacacgccttgaggactatgctgcgaagagtgtgaactccaagaaattgacagtggtgtccgca	889
Db	841	tgaacgcctctggggagtactcgcgcgaagagttaaaacccaatgaattgacggggggcccca	900
OY	890	caaacgcgtgagacatgtgtgtttaaattccgatgtaatccccaaacctcaaccttgcgggtctgac	949
Db	901	caaacgcgtgtgagacatgtgtgtttaatctcgaagccaagcgaagaaccttccacggctctgac	960
OY	950	at'ggaactgaatcatgatgatgaagaataatgaagcccttcggggcaaat--cacagtgctgtca	1006
Db	961	atccaaatgaactctcagaagagatgtatgtgtgtccttcctcgggaacatctgagacaggtgtcga	1020
OY	1007	tgtgtgtcgtcgaagctcgtgtcgtcgtagaagtcttgggttaagtctccgcgaacagcgcaacccc	1066
Db	1021	tggctgtcgtcgaagctcgtgtcgtcgtgagaagtcttgggttaaaatcccgtaaacgacgcacaacct	1080
OY	1067	tatcgtatgtgtgc---taacttaagtgtgggaactgtgtacgaacctgcgggtgtgaacaaacg	1123
Db	1081	gtctcttaagttaacagacagatgaatgtgtgggcactcttaaggagaaactcgcgggtgtgaacaaacg	1140
OY	1124	gaggaaagtcgggtgaatgaacgttcaaatccccaatcctttaaattgtctcaggggcacacacgtgc	1183
Db	1141	gaggaaagtcgggtgaatgaacgttcaaatccccaatcctttaaattgtctcaggggcacacacgtgc	1200
OY	1184	tacaaatgtccgaatacagaaggtctgcgaactcgcgaagagagactaactctctaaagaatctgc	1243
Db	1201	tacaaatgtctcgtatacagaaggtctgtccaaacccgcgaaggttagaacttaactcccaaaaaccca	1260
OY	1244	tcccaagtctggaatgtgggtctgtcgaactcgaaccccaatgaagtcggaaatcgtctgaataatgc	1303
Db	1261	tctgaatctcggaaatcgcgaagctctgcgaactcgcgaatcgcgtgtgaagtcggaaatcgtctgaataatgc	1320
OY	1304	gaaatcagaatctcgcggtatgaatacgttcccggaacactctgtacacacgcgcgttcaacacac	1363
Db	1321	gaaatcagaatctcgcggtatgaatacgttcccggaacactctgtgtacacacgcgcgttcaacacat	1380
OY	1364	cttgaggtgggagaccccggaaggtgtccttctgttaacccgttaagagagacgaactaagaagt	1423

```

Dd      1381 gggagctgggttcaccacaagaatg-ctagtctaacctcggaagcaggtttacacagctg 1439
Oy      1424 aaactcgtaaagggggttgagtcgtatacaacaaggTlac 1459
          |   |||   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1440 tgattcatgactgggggttgaagtcgtaccaagtlagc 1475

RESULT 11
US-09-791-592-1
; Sequence 1, Application US/09791592
; GENERAL INFORMATION:
; APPLICANT: Canon Inc.
; TITLE OF INVENTION: Polyoxyalkanoate containing 3-hydroxybenzoylalkanoic acid
; TITLE OF INVENTION: monomer unit, and method for producing the same.
; FILE REFERENCE: 4386021
; CURRENT APPLICATION NUMBER: US/09/791,592
; CURRENT FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 1
; SEQ ID NO 1
; LENGTH: 1501
; TYPE: DNA
; ORGANISM: Pseudomonas jessenii 161 strain.
US-09-791-592-1
```

Query Match	54.4%	Score 803.2	DB 30	Length 1501
Best Local Similarity	75.9%	Pred. No. 1,2e+247		
Matches 1121	Conservative 0	Mismatches 318	Indels 37	Gaps 9
QY	18 aacgtctgcgcgcgtcttaacaatgcaagtcgacgg-----ggtaagcaata	65		
DB				
QY	3 aacgtctgcgcgcgtcttaacaatgcaagtcgacggtaagtcgtctctctga	62		
DB				
QY	66 ccttaacggcggaaacgggtgagtaacacggtgtatctctccagagtcgtggtaacttc	125		
DB				
QY	63 ttcaacgcgcgcgcgcgcgtgagtaatgacctgaagatctgcgtgaatggggcaacgctc	122		
DB				
QY	126 cgaagcgaaagcaatgaaataacgggaatagtctctgtgtgatacaacaagattgtatagaataat	185		
DB				
QY	123 cgaagcgagcgttaataaccgacata--cgctctacgggaagaagcagggaacctctcggtcc	180		
DB				
QY	166 tatctctggagataagaccgcgcgcgcgtatgaactgtgtgtgagtaagttaagctaccgaag	245		
DB				
QY	181 ttgcgcctatcaatgagccttagcgtgagatctgagatactggtgtgtgagtaagtaagctaccgaag	240		
DB				
QY	246 ggcgaagatcgtgtagcgcgccttgaaagaggtgtctccgcgcacaagaaggaactgagacaggtcc	305		
DB				
QY	241 ggcgaagatcgtgtagcgcgccttgaaagaggtgtctccgcgcacaagaaggaactgagacaggtcc	300		
DB				
QY	306 ataactcttaacggagagcagcagcttaagaaatctgtctcaatgtaggggaacacctgaagcagc	365		
DB				
QY	301 agactctctacggagagcagcagcttgagaaatcttgacaatctggcgaaagacctgataccagc	360		
DB				
QY	366 gacgcgcgcgtgaaacgaagaaggtcttcggatctgtaaaagtctatggcgacgaagaataaag	425		
DB				
QY	361 catgcgcgcgtgtagaagaagaggtcttcggatctgtaaaagactttaaagttggagagaaggg	420		
DB				
QY	426 cagca-----tgtatgatagtgtaactgtgcta--aagcacgcgttaactagc	470		
DB				
QY	421 caataacttaacgttagtgtcttgagccttcacagaaataaagacacgcgttaactctg	480		
DB				
QY	471 tgcacgaacgcgcgttaataacgtatgtgtgcgaacgctgtgttcggaatcatctggcgtaag	530		
DB				
QY	481 tgcacgaacgcgcgttaataacgtatgtgtgcgaacgctgtgttcggaatcatctggcgtaag	540		
DB				
QY	531 ggtgcgtgaagcggatctgtgaatctgaggttgaaactgcygggtctcaaccgtgcactgca	590		
DB				
QY	541 cgcgcgtgaagtggtctgtgaagtgtagatgtgaagcccggtctcaaccctggaactgca	600		
DB				
QY	591 ctgtgaacttaacgaagctctggaagctttggagaagcgaatggaattccaaggtgtagcgttga	650		
DB				
QY	601 ttcaaaactgcaaacgttagatctgtagaagaggtgtgtgaattcttcctgtgtgtagcgttga	660		
DB				



[illegible]

Query Match	54.48;	Score 803.2;	DB 31;	Length 1501;
Best Local Similarity	75.98;	Pred. No. 1.2e-247;		
Matches 1121;	Conservative	0;	Mismatches 318;	Indels 37; Gaps 9;

[illegible]

QY	246	gagagacatcgctagccgagccttggagagggctgctcgcccaaatctggaacttgagacacgtctcc	305
Db	241	ggagacgataccgttaaaccttgctcttggagagatgatatcaatcaacacttggaactcgtgaacacgtctcc	300
QY	306	aatactccataacggggagagacagtttaagaatactctgtctcaatctgggggaaacccctggagacgagc	365
Db	301	agactccataacggggagagacagttgggggaatatattggcaaatctgggcgaagaagccttgatccagc	360
QY	366	gaagccgctgtgaacggaagaagctcttcggatctgtaagtctaatgtctatggcagggaaaaataag	425
Db	361	catgcgcgcgtctgtgaagaagaagctcttcggatctgtaagaagcactttaaatttggaggaaggg	420
QY	426	cagcaac-----ggtagatgatctgactctgcca--aagcacccgcttaactacg	470
Db	421	catcaaaccttaacatcgtaagtgctcttgtaacgtcttaacgacagaataaacacccgcttaactctg	480
QY	471	tgcacagcagccgcgtgtaaatacgtatctgtgcgaagcgtctgtctggaaatctatggcgttaag	530
Db	481	tgcacagcagccgcgtgtatatacagaggctgcaagcgttatactggaattactgtgcgttaag	540
QY	531	ggtctcgtatggcgcggatcttgtaagctcagctgtgtgaataactcgtgcgcttaacccgttgccctgca	590
Db	541	cgcgcgtatggctgtgtctcttgaagtctggaatctggaagaagcccggtgctcaaccttggaactctga	600
QY	591	cttgaatacaaatctccggagctcttgggagagggcaaatctggaaatctccaaagtctagcgtgtga	650
Db	601	ttcaaaactctgaacaaagctaaaggtatggtatagaaggctgtgtgaattctcctgtgtagcgggtga	660
QY	651	atgcgtatagatatcttggaggaacaacacagctggcgaagaagcgactcgtcgtgctcaaaaactgca	710
Db	661	atgcgtatagatataggaaggaaacacacagctggcgaagaagcgactcgtgactgatactgaca	720
QY	711	ctggagcgacggaaagcctgtgggttaagttaacgggatatgaatacccgcttaatccacggctcaaa	770
Db	721	ctgaaagctgtcgaagaagcctgtgggagagcaaacagagatatgaatacccgcttaagctcaacgcgttaa	780
QY	771	acgtgtgtacaaatcttcttggg--ggtcttaaccccttaagtaagaagaaactcaacggtataagt	829
Db	781	acgatgtcaactagccgcttgggagccttggagcctctttagtggtggcgacgtcaacgacttaagt	840
QY	830	agacacgccttgggagactatgctctgcgaagagctgtgaactccaagaagatatgacgggggtctgcga	889
Db	841	tgaacgcctcttgggggtatccggccgaagagtttaaaactcaaatgaattctggccgggggcgcgca	900
QY	880	caagcgcgtggagcacactgggttttaattcgtatgatacccaaaaactccactcgtgggctgtac	949
Db	901	caagcgcgtggagcactgtggttttaattcgtaaagcgaacgcggaagaactcttaacggcctgtgac	960
QY	950	at--ggatctgaaatcaatgttagagabataatgaagcctctcgggcagact--cacaggtgtctga	1006
Db	961	atccaatgtaaactcttcaggaagatatgtatgtgtgacctcttcggagaacatttgaagacaggtctctga	1020
QY	1007	tgtgtgtcgtcaagctctgtctgtgcgtgtgaagatgttgggtttaagctcccgcaacgaagcgaacccc	1066
Db	1021	tgtgtgtcgtcaagctctgtctgtgcgtgtgaagatgttgggtttaagctcccgcaacgaagcgaacccc	1080
QY	1067	tatcgtatctgtgc--tacccttaagtctgggcacacttggtaagaacacttgcgcgtgtgacaaacg	1123
Db	1081	tgtcctttagtttaacgacagcgtlaattgtgtgggcactctaaaggagacatgcccgtgtgacaaacg	1140
QY	1124	gaggggaagcgggtgtatgacatctcaaatctctacatctgctcttaatgtctccaggggccacacacgtgc	1183
Db	1141	gaggggaagcgggtgtatgacatctcaaatctctacatctgctcttaatgtctccaggggccacacacgtgc	1200
QY	1184	tacaaatggccaatcaaaaggtctgcgaacactcgaagaaggaagctaaactcttaaaaagtctgg	1243
Db	1201	tacaaatggctgtacagaaggctgtgcgaagcgcggaggctgtgagctaaatcccaaaaaacgga	1260
QY	1244	tcccaagcttggaattggggctctgtcaactctgaaccccatgaaagctcggaaatcgtatgaatactgc	1303
Db	1261	tctgtatctcggatctgcagctctgtcaaacctcgcagctctgtgaagctcggaaatcgtatgaatactgc	1320







DB 249 TggcgggTaaTggccacc

```

RESULT 1
US-09-967-376-1
: Sequence 1, Application US/09967376
: GENERAL INFORMATION:
: APPLICANT: WALKER, Harrell L.
: APPLICANT: HIGGINBOTHAM, Lawrence R.
: TITLE OF INVENTION: CONTROL OF CYANOBACTERIA WITH A BACTERIUM
: FILE REFERENCE: 013243-0007
: CURRENT APPLICATION NUMBER: US/09/967,376
: CURRENT FILING DATE: 2001-09-28
: NUMBER OF SEQ ID NOS: 1
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 1
: LENGTH: 1540
: TYPE: DNA
: ORGANISM: Unknown Organism
: FEATURE:
: OTHER INFORMATION: Description of Unknown Organism: Bacterium SG-3; gram negative
: OTHER INFORMATION: rod-shaped; exhibits flagellate motility; pathogenic to
: OTHER INFORMATION: cyanobacteria and algae; yellow colonies on BG-11 medium suppl
: US-09-967-376-1

```



QY	283	acaaatggaactcgtgagacacggtccataactctctacccgagagcaacaaatcttaagaatcttgc	342
Db	309	acactctgtgaactctgagacacggtccacagactcttaacgtgagagcaacaaatcttggagaatactgac	368
QY	343	aattggtggaaacccctctgagagcagacacgcccgttgaaacggaagaagctcttcggaattgtaa	402
Db	369	aattggtggcaagctctgatactccagacatggtccgtctgggttgaaagaagcccttcggtttgttaa	428
QY	403	gttcaatgagcagagaaataaagcaag-----caattgcatgatatcttgc	449
Db	429	gcctcttctgtccgaaagaagaagcaactgataataacttcgttctgtcagcgtacccga	488
QY	450	cta--aagacacggctctaaactacgtctccacgacgcccgttgaaatacagatggttgcaagcgt	507
Db	489	agaaataagaccacggtctaaactctgtcccaacgacccgcgtgataatacgaaggtgtcaagcgt	548
QY	508	gtctcgagatacatttggcgtgtaaaggtgtgccttaagcggattcttgaatacgaaggtgtgaaact	567
Db	549	acttcgggaatttaactctggtggtgtaaagcgtgagtgagtggtctgttaaaatctcgtatgtgaaacc	608
QY	568	gcgggtgtcacaacccgttggtccctgtcacctctgaaactacaagtcataagctcttggaagtgcaag	627
Db	609	ctgggtcccaacacttggaattgtcatctgtataactcgtgcgtctagaagtggtcgtgagaagatgc	668
QY	628	ggaaattctcaggtgttaagcgtgtgaaatctgataatctctgaaagaaacacgaatgtgcgaagc	687
Db	669	ggaaatctcccggtgtgataagtgaaatctgcttagagatactggagagaaacactctgtgcgaagc	728
QY	688	gactctgtgtctcaaaactctgacgtctgagacagacaagaacgcttggtgtgtataacggtatga	747
Db	729	gtgcacatctgatacgaagcacttgacacttgaggtcagaaagcgttggtggtgcaaaacagatctga	788
QY	748	taaccggttaatccaacgcccataaagctgtgtctaacagttgtctgggg--tttaaccc	805
Db	789	tacccttgtatgtccacgcccataaagcatgcaactgtgatagttgggtgtcaactaagcactc	848
-QY	806	agtatacgaacacttaagcatttaagttaaacggtcttggtgactatgctcgcgaagatgtgaaact	865
Db	849	agtatacgaacacttaagcatttaagttctgcgcgtcttggtgagttaaagctcgaacactgaaact	908
QY	866	caaaaggaattctgaacggtgtctgcacacaagcgtgtggaacatgtgttcttaattcgaatatacc	925
Db	909	caaaaggaattctgaacggtgtctgcacacaagcgtgtggaatgtgttcttaattcgaatatacc	968
QY	926	ccaaaacactcaaccttggtcttgacatgtgatactgaaat--catgtagaagataatagacct	982
Db	969	cgaagaacacttcaaccttggtcttgacatgtgatacttccagagatgtgatttggctctc	1028
QY	983	tcgggtcgaatatacaagttgtctgatagtgtgtctgtctgaagctcgtgtctgtgagaatgttgggt	1042
Db	1029	gtgggaacacttgacaacagttgtctgatagtgtgtgtctgtcaagctcgtgtctgtgagaatgttgggt	1088
QY	1043	aagtcaccgcgaacagcgaacacccctatctgtattgtc--tacccttaagtttgggcacatg	1099
Db	1089	aagtcaccgcgaacagcgaacaccccttgtcccttaagtttgggcacagcaatgtgtgtgaaactc	1148
QY	1100	taacgaacatgacgtcgtgtgacaacacccggaagagcgtggtatgacgttcaaatctcatgtgct	1159
Db	1149	aggtgagacacgcgtcgtgacaacacccggaagaggttggtgtatgacgttcaaatctcatgtgctc	1208
QY	1160	ttatgtctcaaggtcgaacacacggtgtctataatgtgcgataaaggtgtctgcacactcgcgaag	1219
Db	1209	ttatgtctcaaggtcgaacacacggtgtctataatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	1268
QY	1220	aggtgagcttaactctctataaagctgtgtctccagctcgaattgtggtgtctgcgaactcgcgaacct	1279
Db	1269	aggtgagcttaactctctataaagctgtgtctccagctcgaattgtggtgtctgcgaactcgcgaacct	1328
QY	1280	gaagtctgnaaactcgtctagttaactctgctgagatacga--tgcccgtgtgataagcttcccgagc	1338
Db	1329	gaagtctgnaaactcgtctagttaactctgctgagatacgaatgtgtcgtgtgataagcttcccgagc	1388
QY	1339	ctgtatacacacgcgtctacacacacacttgagttggtgagacacccgaagtgtctctgttatac	1398

[illegible]



```

|||||
Db 473 cgtccagcagccgctaaacgtaggtgctgagcgtgtccggaattactggtgcgtaa 532
QY 529 aggtgtgtaggcggtatttgaatcaggtgtgaaactgtgggtctcaaccgtgtgctg 588
Db 533 agagctcgttaggcgtgttctgctgtctgtgaaaccgcgaactcaactgtcggtctg 592
QY 589 cacttgaactcaacagctctgtaggttggagaagcagtggaattccaggtgtagcggtg 648
Db 593 caggcgatacggcgagacttgaacttgcagaggtgagacttccgtgtgtagcggtg 652
QY 649 aaatgcgtagatatcttgaaggaaacacagctggtcgaagcgagactgtgtccgcaaaactga 708
Db 653 aaatgcgtagatatcaggaagaaacacagctggtcgaagcgagactgtgtccgcaaaactga 712
QY 709 cgtcgtgagcagaaagcgtgtgtgtatgaacgggattagatacccggtatccagcgct 768
Db 713 cgtcgtgagcagaaagcgtgtgtgtatgaacgggattagatacccggtatccagcgct 772
QY 769 aaagctgtctacacaggtgtgtgtgttttaacc---tcagtaacgaaacttaacgatt 825
Db 773 aaacgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 832
QY 826 aaatgaacgcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 885
Db 833 aaagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 892
QY 886 cgtcgtgagcagaaagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 945
Db 893 cgtcgtgagcagaaagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 952
QY 946 tgaatcgtatctgaatcgtatctgaatcgtatctgaatcgtatctgaatcgtatctgaat 1004
Db 953 tgaatcgtatctgaatcgtatctgaatcgtatctgaatcgtatctgaatcgtatctgaat 1012
QY 1005 catgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1064
Db 1013 catgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1072
QY 1065 cctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1121
Db 1073 cctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1132
QY 1122 cgtcgtgagcagaaagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1181
Db 1133 cgtcgtgagcagaaagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1192
QY 1182 gctacaatgtgcgtatcaagaaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1241
Db 1193 gctacaatgtgcgtatcaagaaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1252
QY 1242 ggtccagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1301
Db 1253 ggtccagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1312
QY 1302 ggcgtatcagc-atgcgcggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1360
Db 1313 ggcgtatcagcagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1372
QY 1361 caacttgaactcaacagctctgtaggttggagaagcagtggaattccaggtgtagcggtg 1418
Db 1373 catgaagaatcgttgaacacccgaagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1432
QY 1419 aggttgaactcgttgaagaaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1459
Db 1433 aggttgaactcgttgaagaaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1473

```

RESULT 3  
 US-09-622-201-2  
 : Sequence 2, Application US/09622201  
 : GENERAL INFORMATION:  
 : APPLICANT: MATSUYAMA, Akino bu et al.

```

: TITLE OF INVENTION: Novel microorganisms and processes for producing amide
: TITLE OF INVENTION: compounds
: FILE REFERENCE: 2224-0169P
: CURRENT APPLICATION NUMBER: US/09/622, 201
: CURRENT FILING DATE: 2000-08-14
: PRIOR APPLICATION NUMBER: JP P1998-356229
: PRIOR FILING DATE: 1998-12-15
: NUMBER OF SEQ ID NOS: 2
: SOFTWARE: Patent In Ver. 2.1
: SEQ ID NO 2
: LENGTH: 1480
: TYPE: DNA
: ORGANISM: Unknown Organism
: FEATURE:
: OTHER INFORMATION: Description of Unknown Organism: 16S rRNA gene of
: OTHER INFORMATION: Rhodococcus sp. Am8 (FERM BP-6595)
US-09-622-201-2

```

Query Match 52.9%; Score 780.6; DB 6; Length 1480;  
 Best Local Similarity 75.2%; Pred. No. 2e-259;  
 Matches 1114; Conservative 0; Mismatches 334; Indels 33; Gaps 10;

```

QY 4 catgtcctcagaaactcagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 63
Db 1 cctgtcctcagaaactcagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 60
QY 64 tacc-----tagcgtcgtcagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 110
Db 61 agctgtcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 120
QY 111 tcttggaatacttccgaaagaaagcgtatcagcgtatcagcgtatcagcgtatcagcgtatcagc 170
Db 121 tcttggaatacttccgaaagaaagcgtatcagcgtatcagcgtatcagcgtatcagcgtatcagc 179
QY 171 tgaatgaatgaatcgtatcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 230
Db 180 cgtgtgtgaagaagttactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 239
QY 231 taatgtgtcctcagaaagcgtatcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 290
Db 240 taatgtgtcctcagaaagcgtatcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 299
QY 291 acttggaacagctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 350
Db 300 acttggaacagctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 359
QY 351 aaaccttgaaacagcagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 410
Db 360 aaaccttgaaacagcagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 419
QY 411 ggcgtgaaataaagcagcagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 468
Db 420 agcaggtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 472
QY 469 cgttcgcaacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 528
Db 473 cgttcgcaacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 532
QY 529 aggtgtcgttaggcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 588
Db 533 agagctcgttaggcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 592
QY 589 cacttgaactcaacagctctgtaggttggagaagcagtggaattccaggtgtagcggtg 648
Db 593 caggcgatacggcgagacttgaacttgcagaggtgagacttccgtgtgtagcggtg 652
QY 649 aaatgcgtagatatcgtgagaaacacagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 708
Db 653 aaatgcgtagatatcgtgagaaacacagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 712
QY 709 cgttgaagcagaaagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 768

```



```

QY 838 tggggaactatcgcgaagatgaactcaagaatgacggggtccgcacacggt 897
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 816 tgggaagtaagcgccgaagatgaactcaagaatgacggggtccgcacacggt 875
QY 898 ggaagatggttaactcgaatgatacccaaaaacccacgtgggtcgaatgact 957
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 876 ggaagatggttaactcgaatgatacccaaaaacccacgtgggtcgaatgact 935
QY 958 gaatacgtgagat-----atagagccttcgggacgtatccacgtgctgactg 1012
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 936 gctgtactcgaagatggtatctcccgcaaggaacctccgcacacgtgctgactg 995
QY 1013 tctcaactcgtgtcgtgaatgttggttgaagtcgccgaacgaagcgaacccctacgt 1072
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 996 tctcaactcgtgtcgtgaatgttggttgaagtcgccgaacgaagcgaacccctacgt 1055
QY 1073 atgtgc--taacttaagtgtggactgtgtaacgaactcgcgtgtaacaaacggagaa 1129
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1056 taqtgcgaacgtttgggtgggactcagaagaaactgcgtgtaacaaacggagaa 1115
QY 1130 ggcgggatacgtcaaatcctcatgaccttattgtccaggccacacacgtgtacaat 1189
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1116 ggtgggatacgtcaaatcctcatgaccttattgtccaggccacacacgtgtacaat 1175
QY 1190 ggcggatacgaaggtcgcgaactcgcgaagaaggagatactcttaaaagtcgtccag 1249
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1176 ggcggatgaacgtgggaagcagatgtgtgacacacgtccgacttaaaaag--cgttccag 1234
QY 1250 ttcgattgtgtgtcgaactcgaacccacgaagtcggaatcgtgtaactcgcgtgta 1309
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1235 ttcgattgtgtgtcgaactcgaacccacgaagtcggaatcgtgtaactcgcgtgta 1294
QY 1310 gcatgcgcggtgtgaatacgttcccggaacctgttcaacacgcgcgttaccacacgt 1359
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1295 gcatgcgcggtgtgaatacgttcccggaacctgttcaacacgcgcgttaccacacgt 1354
QY 1370 ggggagacacccgaagtgctctgttgaacgtaaggaacgactataaggtgaactc 1429
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1355 tgggttgaacacggaagcgtg--gtgagcgaacccgcaagacgcaacacacgttgcgtca 1413
QY 1430 gtaaaagggtgtaagtcgttaacaag 1454
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1414 gcgactgtgggtgaagtcgttaacaag 1438

```

```

RESULT 5
US-09-897-516-423
; Sequence 423, Application US/09897516
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Huesling, Joseph E.
; APPLICANT: Krasomil-Osterfeld, Karina C.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spolidonov, Sergei

```

```

; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51847)B
; CURRENT APPLICATION NUMBER: US/09/897,516
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215, 161
; NUMBER OF SEQ ID NOS: 8409
; SEQ ID NO 423
; LENGTH: 83080
; TYPE: DNA
; ORGANISM: Xenorhabdus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (38740)..(39948)
US-09-897-516-423

```

```

Query Match      51.7%; Score 763.4; DB 6; Length 83080;
Best Local Similarity 73.4%; Pred. No. 1,9e-252;
Matches 1100; Conservative 0; Mismatches 356; Indels 43; Gaps 8;

QY 1 gatactgctcaagaactaactcgttcggtcggtcttcaatacgaactcgcggtgtag 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 616 gatcaatgctcagaatgaaactcgttcggtcggtcttcaatacgaactcgcggtgtag 675
QY 61 caatac-----ctagcgcggaacggtgtgagtaaacgttgtaactct 102
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 676 ggaagcagcttgcgtcttgcgttcgaagcgcgcggtgtgagtaaacgttgtaactct 755
QY 103 cctccggtcgtgataactcttcggaagaagcgaatcccgatgactcgtgtgac 162
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 736 cccgattggaaggatataccaactcgtgaacggtgtgctataaccgatacctcgaaga-- 793
QY 163 acaagatttgatagtaaaagattatctgtgtgagatgagccgcggttagtagt 222
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 794 gcaagtggtgggaccttcggtccctacacgtacgtacgtacgtacgtacgtacgt 853
QY 223 tgtgtgagtaactcgtacacgaagcgaacgtgtgtagcgggtcgaaggtgtccggc 202
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 854 aggtggtgtaacggtctacactagcgacgaatccctagctgtgtgagagatgacacgc 913
QY 283 acaatggaactggaacacggtctacactcgtacggtggtggaacggaatctgtctc 342
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 914 acactggtggtggaacacggtccacgactcctacggtggtggaacggtggtggtac 973
QY 343 aatggtggaacacacgtgaaagcgcgcgcgtgtaacgaagaaggtctcgtgataa 402
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 974 aatggtggaacacacgtgaaagcgcgcgcgtgtaacgaagaaggtctcgtgataa 1033
QY 403 gttcatagcaggaata-----aatagcacgaatgtgagatgtact 447
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1034 gttacttcaacggtggaaggaagcgtgtaagtgaatacgttcaacgattgacgtaccgc 1093
QY 448 gccataacacccgtctactacgtgtccacgacgcgcgtgtaatacgttgcgaacgt 507
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1094 agaaagaacacccgtctactacgtgtccacgacgcgcgtgtaatacgttgcgaacgt 1153
QY 508 gttcgaatcattggtcgttaaaagggtgtgtagcgtgattgtaagtcgaagtgtgaaact 567
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1154 aatcggaattactgtgtggttaaaaggcgcgcgcgtgtaatacgttgcgaacgt 1213
QY 568 ggcgtgctcaacccgtgtgcgttgaactgaactgaactgtgtagtagaggtggt 627
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1214 ccgggtcttaacctgtgaactgtcatctgaactgtgtgtagtagaggtggt 1273
QY 628 ggaattcgaagtgtgagcgtgtaaaatgtcgtatatactcgtgaacacacagtggtgaaggc 687
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1274 agaatccacgtgtgtagcgtgtgaaatgtcgtatatactcgtgaacacacagtggtgaaggc 1333
QY 688 gacttgcgtctcaaaactgacgtctgtaggcgaagaacgtgtgtgtagtaaacggtataga 747
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1334 ggcgccctgtgacgaagactgacgtctcaaggtgtgaaagcgtgtgtggaacaagaattaga 1393
QY 748 taacctgtgtagtcaacgcgttaaacgaltgtgactgtggaaggtgtgtggtcctgtgaacgt 1453
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1394 taacctgtgtagtcaacgcgttaaacgaltgtgactgtggaaggtgtgtggtcctgtgaacgt 1453
QY 807 gtaacgaacacgaagatttaagtagacgcgttcgtggtgactatgacgcgaagaagtgaactc 866
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1454 ctccggaactcaacgcgttgaagtagtccgcgttcgtggtgtagtaagtcgcgaagaattaaactc 1513
QY 867 aaagaattgaacggtgtcgcacaaagcgtgtgtagcgtgtgttaacttcgatataacc 926
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1514 aaatgaattgacggtgtggtccgcacacagcgtgtgtagcgtgtgttaacttcgatataacc 1573
QY 927 caaaactcaacccgtgtgacatgtagtcaatcat--gtagagatatagacactcgt 985
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1574 gaagaaccttaactcttgaacatcagcgaatccctcagaagatgaggaaggtcctcgt 1633

```





Oy	1341	gttaacacccgcccgttcaacacacttgaatgggggagacccggaaatggtcttgtttaacg	1400
Db	1994	gttaacacccgcccgttcacacacatggaagatgggtctgaataaagaatag_gtagcttaaacct	2052
Oy	1401	taagaagacagactactaaagtgaagctctgaagaagggttgaaatcgttaacaagttacc	1459
Db	2053	tcggaaaggcgcttacaacacttctgattcaatgaacttgggttgaaatcgttaacaagttaac	2111

## RESULT 8

```

US-09-897-516-3854
Sequence 3854, Application US/09897516
GENERAL INFORMATION:
APPLICANT: Corbin, David R.
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Huesing, Joseph E.
APPLICANT: Krasomil-Osterfeld, Karina C.
APPLICANT: Malvar, Thomas M.
APPLICANT: Slater, Steven C.
APPLICANT: Spiridonov, Sergei
TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
FILE REFERENCE: 38-21(51847)B
CURRENT APPLICATION NUMBER: US/09/897,516
PRIORITY FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: US 60/215, 161
PRIORITY FILING DATE: 2000-06-30
NUMBER OF SEQ ID NOS: 8409
SEQ ID NO 3854
LENGTH: 83080
TYPE: DNA
ORGANISM: Xenorhabdus sp.
FEATURE:
NAME/KEY: CDS
LOCATION: (5826)..(6755)
US-09-897-516-3854

```

Query Match	51.7%	Score 763.4	DB:6	Length 83080
Best Local Similarity	73.4%	Prod. No. 1.9e-252		
Matches 1100	Conservative	0	Mismatches 356	Indels 43
				Gaps 8
QY 1	gacatcgtgctcgaacactacgcctgagcgcgctcgtacttaaacagacagtcacgagcggtag	60		
Db 616	gacatcgtcgtcagatctgaaacgctggtcgcgagcgctcaaacatcagtcacgctgtaaca	675		
QY 61	caatc-----ctagcgcggaacgggtgagtgtaacagctgtaactt	102		
Db 676	ggaagcagctgctgctcttctgctgacgaagcgagcgagctgagatctgctggagatc	735		
QY 103	cctccgagctcgtggaataacttccgaagaagaaagacataacacggaatgacctgttgatc	162		
Db 736	ccggaatgagggggaataccacactggaacacggtgctaatccatgacatcctaagaa--	793		
QY 163	acaagatttgataggttaagaattatttgcgttggaatgagcccgcgccgagatgaactag	222		
Db 794	gcaaaagtggggggaaccttcgagccctccacgcacatcggaatggaacatggaatgaactag	853		
QY 223	tgttgtagtlaatgctcaccagaagcgagatctggttagcgcggtctgaaagagtgctccgsc	282		
Db 854	aggtgggttaacgctcactcactcagcgagcatcccttagctggtctgtagagatgacacagc	913		
QY 283	acaaatggaactgagacagcttccataactctcactcgggagggcagtcagttaaatcttgc	342		
Db 914	acaactggaactgagacacagccagactctcactcgggagggcagtcagttaaatgagc	973		
QY 343	aatgggggaaaccccgaaagcagcgacgacgctggaacgaagaagtgctcggatgttaa	402		
Db 974	aatggcgcaagcccgatgcaagcatgcgcgtgtcatgaaagaagccttcgggtgttaa	1033		
QY 403	gtcatctggcgcaaaa-----aatcagcgaactgtgatagttaact	447		
Db 1034	gtacttcaagcgggaggaagaagcgtgaaatgtataactcgttaccagattpacatccgc	1093		

QY	448	gctcaagcacccgcttaactacgtctgcacgaagccgcggttaactacgtatggttgacaagcgt	507
Db	1094	agaagaagcacccgcttaactccgtctgcacgaagccgcggttaactacgaaggttgacaagcgt	1153
QY	508	gtccgaaactctgtgggttaagaaggtgtcgtaagccggaattcttgaalccaggtgtgaaact	567
Db	1154	aatccgaaattactgtgggttaagaagccggaagccgttcaattgaattgatatgtgaattcc	1213
QY	568	gcggagctcaaccgcgtgcgtgcacttgaaactacaagtcctggaatttggaagagcaagt	627
Db	1214	ccgggactcaacccttggaagactgtcactgaagacttggttgactaagaactctcgttagaaggggggt	1273
QY	628	ggaattcccaagttgtagaggtgtgaaatgtgtagatatcttggaagaaacacacgttgccaaagc	687
Db	1274	aggaattcccaagttgtagaggtgtgaaatgtgtagatatcttggaagaaacacacgttgccaaagc	1333
QY	668	gacttgctgtgtctcaaaccttgacgcgtgagagcacgaagacgcttggttaglaagccggataga	747
Db	1334	ggcccccttgagcagagacttgacgtctcagctgagtgcaagacgcttggtgggacaaacagattaga	1393
QY	748	taccocggttaatccagcccttaaacgtttgtctac-cagtttgtggggttttaaccctca	806
Db	1394	tacccttgtagtaccacccgctlaaacgatatgtctgcacttgtagaggtttgtgccttgtagccgtg	1453
QY	807	gtaaagaaacctcaagggatttaagtagaacccgttggaagactatctccgaagaagtgaattc	866
Db	1454	cttcgcggagctcaaaagcgtttagatgcagccgcctggggagtagcgtctgcagaattaaatactc	1513
QY	867	aaaggaatttgacggyggctccgcacaaagcggctggagacatggtgtttaatctgatatgaacc	926
Db	1514	aaatgaatttgacggyggccgcacaaagcggctggagacatggtgtttaatctgatatgaaccgc	985
QY	927	caaaaactcaacctgggcttgacatgtgatacttgatcat-gttagatatatgagccttcg	985
Db	1574	gaagaaccttcaactcttcttgaaatcccaaggaatcccttcaagagatgtagggaggtgcctcg	1633
QY	986	ggcagatt--cacaggggtgcacatgtgtgtgcgtgcagctcgtctctgtagagatgttggtgta	1043
Db	1634	ggaaagcctgaaagacaggtgctgcgtcaatgtgtctgtcagctcgtgtgttgaaatgttggtgta	1633
QY	1044	agtcocgcgaacgagcgcgaacctcactatcgtatgttc--taacctaaatttgagcactggt	1100
Db	1694	agtcocgcgaacgagcgcgaacctcactatccttctgttcgcacagcagtaagtgttggaactcaa	1753
QY	1101	acgaaactgcgcgggtgacaaacccggaggaagcgggagtagacttcaatctccatgacct	1160
Db	1734	ggggagactgcgcgggtgacaaacccggaggaaggttgaggatgaaacagtaacatcatgacct	1813
QY	1161	tatctccagggcccaacacatgtctatacatgtgcgaatcacagaggtgcgaactccgcaaga	1220
Db	1814	taagagtagggctatacacatgtctatacatgtgcgaatcacagagagtagacctcgcgaga	1873
QY	1221	ggagacttaatctcttaaaagtctgtctccagttcggatttggtgtctgcaactcgaaccacgt	1280
Db	1874	gcaagtcggaactcataaagactctgtctgtagtcggatttgagcttccaactcgaactcact	1933
QY	1281	aagtcggaatgcctgaagtaatccgcggaatcaacatccgcggttaataagttcccggaact	1340
Db	1934	aagtcggaatgcctgaagtaatccgcggaatcaacatccgcggttaataagttcccggaact	1993
QY	1341	gtacacacgcgcgtctacacacactctagttgggagacacccgaaggtgtcttgttatacgt	1400
Db	1994	gtacacacacgcgcgtctacacacatcgggagatgtgtgtgcaaaagaagtag-gtagcttaact	2052
QY	1401	taagagagacagactactaaggtgaactcgtlaaagggtgtgaagtcgtlaacaaggttacc	1459
Db	2053	tcgggagggcgcttaccactttgattcatgactctgggttgaaagtcgtlaacaaggttacc	2111

## RESULT

US-09-897-516-3855  
; Sequence 3855, Application US/09897516



```

: GENERAL INFORMATION:
: APPLICANT: Cordin, David R.
: APPLICANT: Goldman, Barry S.
: APPLICANT: Hinkle, Gregory J.
: APPLICANT: Krasomil-Osterfeld, Karina C.
: APPLICANT: Krasomil-Osterfeld, Karina C.
: APPLICANT: Malvar, Thomas M.
: APPLICANT: Slater, Steven C.
: APPLICANT: Spiridonov, Sergei
: TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences and Uses Thereof
: FILE REFERENCE: 38-21(51847)B
: CURRENT APPLICATION NUMBER: US/09/897,516
: PRIOR FILING DATE: 2001-06-29
: PRIOR APPLICATION NUMBER: US 60/215, 161
: NUMBER OF SEQ ID NOS: 8409
: SEQ ID NO 3855
: LENGTH: 83080
: TYPE: DNA
: ORGANISM: Xenorhabdus sp.
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (7060)..(8658)
: US-09-897-516-3855

Query Match      51.7%; Score 763.4; DB: 6; Length 83080;
Best Local Similarity 73.4%; Pred. No. 1.9e-252;
Matches 1100; Conservative 0; Mismatches 336; Indels 43; Gaps 8;

QY 1 gatcatgctcaagaactaacgcttgccgagcgtcttaaacatgcaagtcgagcgagtag 60
Db 616 gatcatgctcaagaactaacgcttgccgagcgtcttaaacatgcaagtcgagcgagtag 675
QY 61 caatac-----ctagcgagcaagcggtgagtaaacgctgtaactc 102
Db 676 ggaagcagcttgcgtctgtctgtctgacgagcggtgagtagtaatgtctcggagtagctg 735
QY 103 cctccgagcttcggagataactctccgaaagaaagcttaacatgacgtagctctgtgagtc 162
Db 736 ccgagtaggagggagataaacctctgaaacggtggtcttaacacgctgaccccttaagga-- 793
QY 163 acaagatttgatgaagtaaaagtattatgtcttgagatgagtcgagcccgagcagatagtagt 222
Db 794 gcaagtagggagaccttcggcgctcaacgcatcgatgaaccacgagtagtagtagtagt 853
QY 223 tggcgagtagtagctcaacgagcagatcgtagcggcgccgagaggggtgtccggc 282
Db 854 aggtggggttaacggtctccctcctagcgagatcccttagctgtgtctgagagtagtaaccagcc 913
QY 283 acaatgagactgagacacggtccatactcctcagggagcgagcagcttaagaactctgtctc 342
Db 914 acaatgagactgagacacggtccatactcctcagggagcgagcagctgagagatattgac 973
QY 343 aatgggggaaacccctgaaagcagcgagcggtgaaagaaaggtcttcgagtagttaa 402
Db 974 aatggcgcaagcctgtagtcagcagctgagcggtgtatgaagaagcgcttcggtttgtaa 1033
QY 403 gtctcattagcaggaagaa-----ataagcagcaatgtgtagtctact 447
Db 1034 gtactctcagcgaggaggaagcgtagaagttgaatagcttccacatgtgacgttaccgc 1093
QY 448 gccataagcagcgctaaactacgtgcacgagcgcgtaatacgtatgctgaagcgt 507
Db 1094 agaagaagcagcgctaaactacgtgcacgagcgcggttaatacgtgaagcgtagcagcgt 1153
QY 508 gtccgagacatcgtggtgctaaaggtgctgtagcgagattgtgaagtcaggtggaacact 567
Db 1154 aatcggaattactggcgtagaagcgacgagcggttcaattaaagtagatgtgaataacc 1213
QY 568 gcgaggtcaaacgcttggtcctcacttgaaactaaacagcttgaggttgggaggaagcagt 627
Db 1214 cggaggtcaaacctggagactgtacataagactgttgaactagagctcgttagaagggtt 1273
```

```

QY 628 ggaattccaggttagcggtgaatgctgtagatattcggaagaaacacagctggcgaagc 687
Db 1274 agaattccaggttagcggtgaatgctgtagatattcggaagaaacacagctggcgaagc 1333
QY 688 gactgtcgtgctcaaaaactgacgctgtagagcgaaagcggtgtgtagtaaacggagtaga 747
Db 1334 ggcctccgagcaagagactgacgctgtagagcggtgtgtagtaaacggagtaga 1393
QY 748 taaccggttatccacgcccataaagttgtctac- caagtgttgggggttttaaccctca 806
Db 1394 taaccgttagtccacgcccataaagttgtctac- caagtgttgggggttttaaccctcg 1453
QY 807 gtaagaaactaaacggatagtagtagcgcctgggagtagtagcttcgcaagaagtgaaactc 866
Db 1454 ctcccgaggtcaacggttagtagtagcgcctgggagtagtagcttcgcaagaagtgaaactc 1513
QY 867 aaaggaattgacgggggtcgcacaaagcggtgagacatgtgtttaattcgtatgataacc 926
Db 1514 aatgtgaattgacgggggtcgcacaaagcggtgagacatgtgtttaattcgtatgataacc 1573
QY 927 caaaacctcaacctggtgcttgcacatgtagtgaatcat- gtagagatatatgaccttcg 985
Db 1574 gaagaaacctcaaccttcttgcacatgtagtgaatcat- gtagagatatatgaccttcg 1633
QY 986 ggcagatt--cacaagtgctcatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1043
Db 1634 ggaagcgtgagacaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1693
QY 1044 agtcccgcaacgagcgcaacccctatcgtatgttcg---taacctaaatttgggcaactgt 1100
Db 1694 agtcccgcaacgagcgcaacccctatcgtatgttcg---taacctaaatttgggcaacttaa 1753
QY 1101 acgaactgtccggtgacaaacccgaggaagcggtgagtagcgttcaactcctatgacct 1160
Db 1754 gggagactgtccggtgagaaacccgaggaagcggtgagtagcgttcaactcctatgacct 1813
QY 1161 tatgtccacgccaacacggtgtctacatggtcgatagagaggggtcgcgaactcgcgaaga 1220
Db 1814 taagtagtggtgtacacacggtgtctacatggtcgatagagaggggtcgcgaactcgcgaaga 1873
QY 1221 gggagactatcttaaaagctgtgtccagcttcgagattgggggtctgcaactcgcgaactg 1280
Db 1874 gcaagcgtgaactcataaagctgtgtccagcttcgagattgggggtctgcaactcgcgaactg 1933
QY 1281 aagtcggaatcgtctgtagtgcgagtagcagatcgcggtggaatagcttcccgacct 1340
Db 1934 aagtcggaatcgtctgtagtgcgagtagcagatcgcggtggaatagcttcccgacct 1993
QY 1341 gtacacacgcccgttcacacacacccctgagtggtggaacccggaagtgtgtctgttaacg 1400
Db 1994 gtacacacgcccgttcacacacacccctgagtggtggaacccggaagtgtgtctgttaacg 2052
QY 1401 taagtagacagactactaaagttgaaactcgtaaaggggtgaagtctgtaacaagtagc 1459
Db 2053 tcgggagggcgcttaacacttgtgtattcatgactgtgggtggaagtcgtaacaagtagc 2111

RESULT 10
US-09-897-516-3856
: Sequence 3856, Application US/09897516
: GENERAL INFORMATION:
: APPLICANT: Cordin, David R.
: APPLICANT: Goldman, Barry S.
: APPLICANT: Hinkle, Gregory J.
: APPLICANT: Krasomil-Osterfeld, Karina C.
: APPLICANT: Krasomil-Osterfeld, Karina C.
: APPLICANT: Malvar, Thomas M.
: APPLICANT: Slater, Steven C.
: APPLICANT: Spiridonov, Sergei
: TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences and Uses Thereof
: FILE REFERENCE: 38-21(51847)B
: CURRENT APPLICATION NUMBER: US/09/897,516
```

```

?
? CURRENT FILING DATE: 2001-06-29
? PRIOR APPLICATION NUMBER: US 60/215, 161
? PRIOR FILING DATE: 2000-06-30
? NUMBER OF SEQ ID NOS: 8409
? SEQ ID NO 3856
? LENGTH: 83080
? TYPE: DNA
? ORGANISM: Xenorhabdus sp.
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (8720)..(10027)
? OS=09-897-516-3856

```

Query Match	51.7%	Score 763.4	DB 6	Length 83080
Best Local Similarity	73.4%	Pred. No. 1.9e-252		
Matches 1100; Conservative	0	Mismatches 356	Indels 43	Gaps 8

QY	1	gatactgacccaagaacaaacgctgagcgcgctcttaacacatgcaagctcggaacggtgac	60
Db	616	gatactgacccaagaatggaacgctgagcgcgacgacctaacacatgcaagctcggaacggtgac	675
QY	61	caatac-----ctagcggcggaacggtgagtaacacgttgatactt	102
Db	676	ggaagcagctctgctgtcttgctgacgaacggtcggaacggtgagtaatgtcttggtgactg	735
QY	103	ccctccgagctctggatgaactcttcggagaaggaagacataacccggatgacctgttgatc	162
Db	736	cccgatgtaggggataaaccaactctggaacacgtgagctaaataacgcatagaccttaagga--	793
QY	163	acaagattgatagtgataaagaatttatgtcttgagatatgagcccgcgccgaattgaatagt	222
Db	794	gcaaaagtggggagacctctcggtcctcaacgcatcagatgagtaaccagatggtatgactagt	853
QY	223	tggctgtagatattggtcccaacaagcggaagatctgtagcccgccctgaagaggtgtccggcc	282
Db	854	aggtctgggtgataacggctccacctaaaggcaacatcccttaagctggtctctgagagatagac	913
QY	283	acaaatggaactgagacaacggtccatcactctcaacggagagcaacaaatgaagactctgttc	342
Db	914	acactctggaactgagacaacggtccgaagctctctcaacggagagcaacaaatgaagactctgttc	973
QY	343	aattgggggaaacccctgaagcagcgacgcccgttgaaacgaagaagctctcgattgtgtaa	402
Db	974	aattggggcgaacacgtatgatacgaatgcccgtgtgataagaagacgtcttcggtgtgtaa	1033
QY	403	gttccattaggccgaaga-----aataagacgaatgtagatgtagtact	447
Db	1034	gttccattcaacgggggtaggaagcgtgtgaagtctgaatacgttcaacatgtaacgtttaacccg	1093
QY	448	gctctaaagcaccggctcaactacgtctgcacaacgacgcgcgtgaatacgtatgtagtcaagcgt	507
Db	1094	agaagaagacaccggctcaactccgtgcacagcagccgcgtgaatacgaaggtgtgcaagcgt	1153
QY	508	gttcggaatcatctgggctgataaaggctgcgttgagcggattgtgaaatcaggttgtaaaact	567
Db	1154	aattcggaaatcatctgggctgataaaggcgcacgacgacggttcaatgaattagatgtgaatacc	1213
QY	568	ggggggtccaacccgttgccctgcacatggaactacaagaatcttgagaatttgtaggaaggaacgt	627
Db	1214	ccggggttcaacccgttggaactgcgtcccaagaactggttgactaaagttctcgttgaagggggtc	1273
QY	628	ggaattccaaagtgtgaacggttgaaatgctgataatcttgtaggaagaacacgaatgtgcgaagc	687
Db	1274	agaattccaaacgtgtagcgtgtgaaatgctgataagatgttgtaggaataacacggtgtgcgaagc	1333
QY	688	gaattctgtgctcaaaactggaacgttgaaagcgaacgaacgctgttgtagtaacggaatga	747
Db	1334	ggcccccttgagaagaagatcgtacgcttaagctggaagaacgctgtgggggacaaacagaatga	1393
QY	748	taccctggataacacgcacctaaacgtgtcttacc--cagttgttggtgttttaacctca	806
Db	1394	taccctggataacacgcacctaaacgaatctgcgaactgttgaggttggtgttgaccttgagccgtg	1453

Oy	807	gtaacgaacctaaacggatttaagttagacgcctctggggactatgtctcgaagaaatgaacac	866
Db	1454	cttcgcgaagctcctaaacgcgttaagtctgaacgcgcctggggagtaacgcgtcaataaatac	1511
Oy	867	aagaagatltgaacggygggtlccgcacaaacggttgagacatgtgttlaattcgaatgaaccc	926
Db	1514	aattgaattgacggygggtccgcacaaacggttgagacatgtgttlaattcgaatgaacgc	1572
Oy	927	caaaaactcaacctgggctcttgacatgtgattctgaatcaat-gtagaataatagacctcg	985
Db	1574	gaagaaccttcaacctctcttgacatctcacaagaaatccctcaagaatltagggagtgacctcg	1633
Oy	986	ggccgatt--cacaggtgctctgacatgtgtgtcgtcaactcgtgtgcgtgaagtcttggtta	1045
Db	1654	ggaaacgctgagaacgtgcgtgcgaatgagctgtcgtcaagctcgtgtgtgaaatgttggtta	1693
Oy	1044	agtcctcgcaacgaagcgacgaacccctatcgtlactgtgc---tacctaaagtltggcacgtgt	1100
Db	1694	agtcctcgcaacgaagcgacgaacccctatccttctgttctgcagaacgtaaatgttggaactcaa	1755
Oy	1101	acgaacatctgcgtgtgacaaacccggaagagcgggagctgaacgtlcaaatccctcaatgcctt	1160
Db	1754	gggagacatctgcgtgtgacaaacccggaagagctgtggagctgaacgtlcaaatccctcaatgcctt	1813
Oy	1161	tatgtccgaagggccacacagctgtctacaaatgtgcgatacagaaggtgtgcacactctgcaaga	1220
Db	1814	tacgaagctagggctctacacacgtgtctacaaatgtgcgatacagaagaggaagacactctgcgaga	1873
Oy	1221	gggagctaatctctcaaaagctcgtgtccacgtctcgatttgggtctgtcgaactcgaacccatg	1280
Db	1874	gcaagcggaactctcaaaagctctgtgttcgttcgcgatttgagctgtcgaactcgaacccatg	1933
Oy	1281	aagtcggaatctgcgtgaatcgcgagatcacacatgtccggttgaaatagcttcccgaaacct	1340
Db	1934	aagtcggaatctgcgtgaatcgcgagatcacacatgtccggttgaaatagcttcccgaaacct	1993
Oy	1341	gtacacacgcgcgcgtctcaacacaccttgatgtggagagcacccggaagctgtcttctgttaacgg	1400
Db	1994	gtacacacgcgcgcgtctcaacacaccttgatgtggagagcacccggaagctgtcttctgttaacgg	2055
Oy	1401	taagagagacagactactcaaggtgtgaactcgttaaggggggttgaaatcgttaacaaggtaac	1459
Db	2053	tcgggagggcgcttaccactctgtgattctacgtacttgggttgaaatcgttaacaaggttaac	2111
RESULT 11			
US-09-897-516-3857			
Sequence 3857, Application US/09897516			
GENERAL INFORMATION:			
APPLICANT: Corbin, David R.			
APPLICANT: Goldman, Barry S.			
APPLICANT: Hinkle, Gregory J.			
APPLICANT: Huesling, Joseph E.			
APPLICANT: Krasomil-Osterfeld, Karina C.			
APPLICANT: Malvar, Thomas M.			
APPLICANT: Slater, Steven C.			
APPLICANT: Spiridonov, Sergei			
TITLE OR INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof			
FILE REFERENCE: 38-21(51847)B			
CURRENT APPLICATION NUMBER: US/09/897,516			
CURRENT FILING DATE: 2001-06-29			
PRIOR APPLICATION NUMBER: US 60/215, 161			
PRIOR FILING DATE: 2000-06-30			
NUMBER OF SEQ ID NOS: 8409			
SEQ ID NO 3857			
. LENGTH: 83080			
. TYPE: DNA			
. ORGANISM: Xenorhabdus sp.			
FEATURE:			
NAME/KEY: CDS			
LOCATION: (10097)..(11929)			
US-09-897-516-3857			



Query Match 51.7%: Score 763.4; DB 6: Length 83080;  
Best Local Similarity 73.4%: Pred. No. 1.9e-252;  
Matches 1100; Conservative 0; Mismatches 356; Indels 43; Gaps 8;

QY 1 gatcatggtcagaactacgctggtcggtgcttaacacatgcgaagtcggtgtag 60  
Db 616 gatcatggtcagaactacgctggtcggtgcttaacacatgcgaagtcggtgtag 675

QY 61 caatac-----ctagcgcgaaacgggtgagtaacacgtgtaactct 102  
Db 676 ggaagagctgtctgtctgtctgtcgaagcggtgtagtaactctcgtggaactctg 735

QY 103 cctcgaagctcgtggaactctcgaaggaagtaactacggaatgctcgtgtgact 162  
Db 736 ccgatgtgagggggaataacacacgtggtgtaactacggaactggaactcctaagg-- 793

QY 163 acaagatctgaatgaatgaatctatctgtgagatgagccggtcggtgtagtagt 222  
Db 794 gcaagatggtggaactctcgtcgaacgtcgaacgtcgaacgtggaatgtagtagt 853

QY 223 tgggtgagtagtgcgtcgaacgtcgaacgtcgtgtagcgtcggtgtagtagt 282  
Db 854 aggtggtgtagcgtcgaacgtcgaacgtcgtgtagcgtcggtgtagtagt 913

QY 283 acaatggaactggaacacgtcgaacgtcgaacgtcgaacgtcgaacgtcgaacgt 342  
Db 914 acactgtggaactggaacacgtcgaacgtcgaacgtcgaacgtcgaacgtcgaacgt 973

QY 343 aatgtggtggaacacgtcgaacgtcgaacgtcgaacgtcgaacgtcgaacgt 402  
Db 974 aatgtggtggaacacgtcgaacgtcgaacgtcgaacgtcgaacgtcgaacgt 1033

QY 403 gttcaatgagcaggaacacgtcgaacgtcgaacgtcgaacgtcgaacgt 447  
Db 1034 gttcaatgagcaggaacacgtcgaacgtcgaacgtcgaacgtcgaacgt 1093

QY 448 gctcaatgagcaggaacacgtcgaacgtcgaacgtcgaacgtcgaacgt 507  
Db 1094 agaaagaaacacgtcgaacgtcgaacgtcgaacgtcgaacgtcgaacgt 1153

QY 508 gttcgaatcgtggtcgaacgtcgaacgtcgaacgtcgaacgtcgaacgt 567  
Db 1154 aatcgaatcgtggtcgaacgtcgaacgtcgaacgtcgaacgtcgaacgt 1213

QY 568 gctggtcgaacacgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 627  
Db 1214 cgtggtcgaacacgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1273

QY 628 ggaatcgaacgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 687  
Db 1274 agaatcgaacgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1333

QY 688 gaattggtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 747  
Db 1334 ggcctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1393

QY 748 taaccggtgaatcgaacgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 806  
Db 1394 taaccggtgaatcgaacgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1453

QY 807 gtaacgaacacacgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 866  
Db 1454 ctccggaacacacgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1513

QY 867 aaaggaatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 926  
Db 1514 aaaggaatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1573

QY 927 caaaacacacacgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 985  
Db 1574 gaagaaacacacgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1633

QY 986 ggcagatc--cacaggtgctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1043  
Db 1634 ggaacgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1693

QY 1044 agtcccgaaacacacgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1100  
Db 1694 agtcccgaaacacacgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1753

QY 1101 agaaactcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1160  
Db 1754 ggaagactcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1813

QY 1161 tatgtccgggacacacgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1220  
Db 1814 taagatggtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1873

QY 1221 ggaagtaactcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1280  
Db 1874 gcaagcgaactcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1933

QY 1281 aagtcgaatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1340  
Db 1934 aagtcgaatcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1993

QY 1341 gtacacacgcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1400  
Db 1994 gtacacacgcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 2052

QY 1401 taaggaagacacacgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 1459  
Db 2053 tcgggaggtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 2111

RESULT 12  
US-09-897-516-3858  
Sequence 3858, Application US/09897516  
GENERAL INFORMATION:  
APPLICANT: Corbin, David R.  
APPLICANT: Goldman, Barry S.  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Huesling, Joseph E.  
APPLICANT: Krasomil-Osterfeld, Karina C.  
APPLICANT: Malvar, Thomas M.  
APPLICANT: Slater, Steven C.  
APPLICANT: Spiridonov, Sergei  
TITLE OR INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof  
FILE REFERENCE: 38-21(51847)B  
CURRENT APPLICATION NUMBER: US/09/897,516  
CURRENT FILING DATE: 2001-06-29  
PRIOR APPLICATION NUMBER: US 60/215, 161  
NUMBER OF SEQ ID NOS: 8409  
SEQ ID NO 3858  
LENGTH: 83080  
TYPE: DNA  
ORGANISM: Xenorhabdus sp.  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (17551)..(19197)  
US-09-897-516-3858

Query Match 51.7%: Score 763.4; DB 6: Length 83080;  
Best Local Similarity 73.4%: Pred. No. 1.9e-252;  
Matches 1100; Conservative 0; Mismatches 356; Indels 43; Gaps 8;

QY 1 gatcatggtcagaactacgctggtcggtgcttaacacatgcgaagtcggtgtag 60  
Db 616 gatcatggtcagaactacgctggtcggtgcttaacacatgcgaagtcggtgtag 675

QY 61 caatac-----ctagcgcgaaacgggtgagtaacacgtgtaactct 102

```
Db 676 ggaagcagctctgctcttctgtctgacgagcgcgagcggtgtgaattgtctgggtctctg 735
Qy 103 cctccgagctctggataactctccgaaagaaagctaatccggaatgcctgtctgacc 162
Db 736 ccgagatggagggtgataaaccaacttggaacggtgtgctaatccgcatgaccttaagaa-- 793
Qy 163 acaagaatttgataagtaaaagatttactctgtgagaatgacgcgcgcgcgcgcgcgcgcgcgc 222
Db 794 gcaaaagtgggggacactctcggtccacgcacatcgtgacatgaaacccagatgggattagctagt 853
Qy 223 tggtagtgtaatgtgctcaccacaagcgagatcgtagccggtcgtgagaaggtgtccggcc 282
Db 854 aggttggtgtaacggtcgcacacttagcgacatcccttagctgtgtctgagaaagtatgacacgc 913
Qy 283 acaatggaacttgagacacgcttccatctccttactagaggagcagcagttgaagaatcttgctc 342
Db 914 acaatgggacttgagacacgctccacacgtcctcctacggaagcagcaggtggggaataatgtcac 973
Qy 343 aatgggggaaacccctgaagcagcgacgcgcgtgaaacgaagaagttctcgtgaattgtaaa 402
Db 974 aatgggcgcaagcctgtatgcagccatgcgcggtgtatgaaagaagccttcgggtctgtaaa 1033
Qy 403 gtctcataggcagga-----aataagacgaatgtgatatgtaacct 447
Db 1034 gtaacttcagcggggaggaagcggtgaagttgaaatacgttcttcgacatgacgttaccgcgc 1093
Qy 448 gcttaaagcaccggtctaactacgttgccaagcgcgcgtgaataacgtatgtgcaagcgtct 507
Db 1094 agaagaagcaccggtctaactcgttgcgcagcgcgcgtgtaatacgaagggttgaagcgtct 1153
Qy 508 gtctcggaatcatttgggcgtgaaggggtgcgttaggcggaatttgttaagtcaagttgtgaaact 567
Db 1154 aatcggaattacttgggcgtgaagcgcgacacgagcggttcaatlaagttagatgtgaaatcc 1213
Qy 568 ggaggcgtaacccggtgcgtgcacttgaaactgaacgaatctggaagtttgggaagggaagct 627
Db 1214 ccggggttaactcttggaactgtcaactgaagctgtgtgactagaagctctccgaaggggggtct 1273
Qy 628 ggaatccaggtgtctagcgggtgaaatgtctagataatcttggaaggaaacccagttggcggaagc 687
Db 1274 agaatctcagcgtgtagcgggtgaaatgtcttagagatgttggaagataccggttgcggaagc 1333
Qy 688 gaactgtgctgcataaacctgtagcgtctgaggcaacgaagcgttgggttagtaaacgggattaga 747
Db 1334 ggcaccttgtagaagaagcgtgaagctgtaggtgtggaagcgttggggaacaaacggattaga 1393
Qy 748 taaccggtgaatccacgaccttaacgttcttac--cagttgttgggggtttaaaccctca 806
Db 1394 taacctgtgtagtcacacgcgttaaacgatagttcgacttggaggttctgtgccttgagccggtg 1453
Qy 807 gtaacgaacctaagatgaatgaacgcgtctggggatactgctcgcgaagatgaaactc 866
Db 1454 ctctccgagctaagcgttgaatgaatgcacgcgtcgggagtagtcgcgaagatlaaaactc 1513
Qy 867 aagaagaattgaaagggtctgcgaacaaagcggtggaatgtgtgttaattcgaatgataccc 926
Db 1514 aaaaatgaatggaggggcccgcgaacaaagcggtggaatgtgtgttaattcgaatgataccc 1573
Qy 927 caaaaactcaccctgggcttgatgaatgatactgatacat--gttagagaatatagaccttcg 985
Db 1574 gaaagaaccttaactactctcttgacatccagaaatcctctcaagagatgaaggagtgccttcg 1633
Qy 986 ggcagatt--caacaggtgtgtgacatgtgtctgctcagctcgtgtctgtgaatgttgggtta 1043
Db 1634 gaaacgctgagacaggtgtgtgtgacatgtgtctgctcagctcgtgtgtgtgaatgttgggtta 1693
Qy 1044 agtccggaagagcgcaacccctacgtatgttgc---taacctaaagtctggcactgtgt 1100
Db 1694 agtcccgcaacagcgcaacccctacgtcttctgtccacagctaaatgtgttgggaactcaaa 1753
Qy 1101 acaaaactcgcgttgacaacacgggaagcggtggaatgacgttcaactcctcaatgtgcctt 1160
Db 1754 gggagagctgcggtgtgataaacgggaagaggtgtgggatagtacgtcaatcataatgtgcctt 1813
```

```
Qy 1161 tatgtccagggccacacagctgtctataatggtccgatcacagaggtgtccgaactcgcgaaga 1220
Db 1814 taagagtaagggtccacacagctgtctataatggtcagatacaaaagaagcgacctcgcgaga 1873
Qy 1221 gggagctaatctctaaagtlcgtgtccagcttcggaattgggtgtctcgaactcgaacccatg 1280
Db 1874 gcaagcggaactcataaagtlcgtgtccagcttcggaattgggtgtctcgaactcgaactccatg 1933
Qy 1281 aagtggaaatcgttagtaatcgcggaatcagcatgcgcggtgtgaatacgttcccggaactt 1340
Db 1934 aagtcggaatcgttagtaatcgcggaatcagcatgcggtgtgaatacgttcccggaactt 1993
Qy 1341 gtaacacggcccggtcacacccactgagttggggagaccccggaaggtgtctctgtgtaacgg 1400
Db 1994 gtaacacggcccggtcacacccactgagttgggtgtgaaagaagatgtg--gttagcttaact 2052
Qy 1401 taaggagacagactactaaagtgtaaactcgttaaaagggtgtgaagtctgtaacaagtacc 1459
Db 2053 tcggagagcggtccttaaccacttgtgtatcatgacttgggtgtgaagtcgttaacaagtacc 2111
```

```
RESULT 13
US-09-897-516-3859
Sequence 3859, Application US/09897516
GENERAL INFORMATION:
APPLICANT: Corbin, David R.
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Huesing, Joseph E.
APPLICANT: Krasomil-Osterfeld, Karlina C.
APPLICANT: Malvar, Thomas M.
APPLICANT: Slater, Steven C.
APPLICANT: Spiridonov, Sergei
TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
FILE REFERENCE: 38-21(51847)B
CURRENT APPLICATION NUMBER: US/09/897,516
CURRENT FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: US 60/215, 161
PRIOR FILING DATE: 2000-06-30
NUMBER OF SEQ ID NOS: 8409
SEQ ID NO 3859
LENGTH: 83080
TYPE: DNA
ORGANISM: Xenorhabdus sp.
FEATURE:
NAME/KEY: CDS
LOCATION: (37881)..(38108)
US-09-897-516-3859
```

```
Query Match 51.7%; Score 763.4; DB 6; Length 83080;
Best Local Similarity 73.4%; Pred. No. 1.9e-252;
Matches 1100; Conservative 0; Mismatches 356; Indels 43; Gaps 8;

Qy 1 gatcatgtctcgaactaacgcgtgcgcgcgtcttaaacatgcaagtcgagcggtgtag 60
Db 616 gatcatgtctcgaactaacgcgtgcgcgcgtcttaaacatgcaagtcgagcggtgtag 675
Qy 61 catatc-----ctagcggcgaaaggggtgtgaacacgtgtgtaactctt 102
Db 676 ggaagcagctctgcttctgtctgtagcagcgcggaagggtgtgaaatgtctgtgggtctg 735
Qy 103 cctccgagctctggataactctccgaaagaaagctaatccggaatgcctgtctgacc 162
Db 736 ccgagatggagggtgataaaccaacttggaacggtgtgctaatccgcatgaccttaagaa-- 793
Qy 163 acaagaatttgataagtaaaagatttactctgtgagaatgacgcgcgcgcgcgcgcgcgcgcgc 222
Db 794 gcaaaagtgggggacactctcggtccacgcacatcgtgacatgaaacccagatgggattagctagt 853
Qy 223 tggtagtgtaatgtgctcaccacaagcgagatcgtagccggtcgtgagaaggtgtccggcc 282
```

```
Db 854 agtggggttaacggctcacctagagcgacgataccctagctgtgtctgtagaggaatgacagcc 913
Oy 283 acaatggagactggagacacgagctccatctctacgggaagcgagcaatlaagatctgtctc 342
Db 914 acaatggagactggagacacgagctccatctctacgggaagcgagcaatlaagatctgtctc 973
Oy 343 aatggggaacacccctgaagcgagcgagcgagctgaacgaagaaggtctctcgatgtgtaa 402
Db 974 aatggggaacacccctgaagcgagcgagcgagctgaacgaagaaggtctctcgatgtgtaa 1033
Oy 403 gttaactagcgagga-----aataagcagcaatgtatgtactc 447
Db 1034 gtactctagcgagggaaggaagcgctgaagtgtgaatacttcaacgattgacgttaccgcg 1093
Oy 448 gacctaaagcagcggttaactctgtccagcagcgccggtaactacgtatggcgcaagcgct 507
Db 1094 agaaagaacacgagcttaactccgtgacagcagcgaggtaatacggaggtgcaagcgct 1153
Oy 508 gttcggaatcattggcgtaaaaggtgtagcgagattgtlaagtcaggtgtagaaact 567
Db 1154 aatcggaattactgtggcgtaaaagcgacagcgagcggttaactaagtgtagtaaatcc 1213
Oy 568 gcgaggtcaacccgtgagctgagctgaactgaactaagctgtaggttggggaagcgcaagt 627
Db 1214 ccgggcttaactgtggaactgcatctaaagactggttgaactagagctcgttagaagggggt 1273
Oy 628 ggaattccaggtgttagcggtgaatgcgttagatattctggaagaacacgaaggcggaagc 687
Db 1274 agaatccacgtgtagcggtgtaaaatgctgaagatgctggaagaaataccgggtgcaaaagc 1333
Oy 688 gactgtcgctcaaaaactgaacgtgagcaagaagcgtaggtgaatgaacgggattaga 747
Db 1334 ggcgccctggaggaagactgagcaggtgcaagagcgtaggggagcaaacgggttaga 1393
Oy 748 taaccggtaatacaccgcttaaacggtgtgtac--cagttgtgggggtttaaaccctca 806
Db 1394 taaccgtgttagtccacgcgttaaaacgattgctgaacttggagcttggccttgaagcgtg 1453
Oy 807 gtaaggaactcaacgagattagtagccgctggggaactatgctgcgaagagtgaaactc 866
Db 1454 ctcccgaggtctaaacggttaagatgacccgctggggaatgaagtcgcgaagataaaactc 1513
Oy 867 aaaaagaattgacgggggtgcgcacaaagcggtggaagcatgtgtttaaactgaatgaacc 926
Db 1514 aaatgaattgacgggggtgcgcacaaagcggtggaagcatgtgtttaaactgaatgaacgc 1573
Oy 927 caaaaactcaactgtggttgaatgacatgatactgaaatcat--gtagaatatatgaactctc 985
Db 1574 gaaagaactcaactgtggttgaatgacatgatactgaaatcatgaaaggaatgtgctctcg 1633
Oy 986 ggaagatt--cacaaggtgctgcatgtgtgtcgtacgctggtgctgctgtagatgttggtta 1043
Db 1634 ggaagctggaagacaggtgctgcatgtgtgtcgtacgctggtgctggttgaatgttggtta 1693
Oy 1044 agtcccggaacgagcgcaaccctatcgtaatgtgc---taacttaagtltgggcaactggt 1100
Db 1694 agtcccggaacgagcgcaaccctatcgtaatgtgtcgtacgcaacgtaatgtgtggaaactcaa 1753
Oy 1101 acgaagactgcccgtgtaacaaacggggaagcggtgtagatgcaaatccctatgtgctt 1160
Db 1754 ggggaactgcccgtgtaacaaacggggaagcggtgtagatgcaaatccctatgtgctt 1813
Oy 1161 tatgcccagggccacacacgtgtctacaaatgtagcgaatacgaaggtgcgcaactgcgaaga 1220
Db 1814 taaggtatgggtgtataacacgtgtctacaaatgtagcgaatacgaaggtgcgcaactgcgaga 1873
Oy 1221 gggagataactctaaagctggttccagcttcggaattgggtctgcaactcgaccccatg 1280
Db 1874 gcaacggaactcaataaagctgtctgtagtccggaattgtagtctgcaactcgaccccatg 1933
Oy 1281 aagtcggaatcgttagtaactgtagcgaatacgaatgtagcgggtgaataaggttcccggaact 1340
Db 1934 aagtcggaatcgttagtaactgtagcgaatgtagcgggtgaataaggttcccggaact 1993
```

```
Oy 1341 gtacacacgcccgtacacaccacccctgaatggggagacccgaagtgtcttgttaacg 1400
Db 1994 gtacacacgcccgtacacaccacccctgaatggggagacccgaagtgtcttgttaacact 2052
Oy 1401 taaggaagacactaactaagtgtaaacctcgtlaaagggggtgaagtgtgtaaacggtacc 1459
Db 2053 tcgggaaggcgcttaccacttgtgtatcatgactggtgggtggaagtcgtaacgtaacgtaac 2111

RESULT 14
US-09-897-516-3860
: Sequence 3860, Application US/09897516
: GENERAL INFORMATION:
: APPLICANT: Corbin, David R.
: APPLICANT: Goldman, Barry S.
: APPLICANT: Hinkle, Gregory J.
: APPLICANT: Huesling, Joseph E.
: APPLICANT: Krasomil-Osterfeld, Karina C.
: APPLICANT: Malvar, Thomas M.
: APPLICANT: Slater, Steven C.
: APPLICANT: Spiridonov, Sergei
: TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
: FILE REFERENCE: 38-21(51847)B
: CURRENT APPLICATION NUMBER: US/09/897,516
: CURRENT FILING DATE: 2001-06-29
: PRIOR APPLICATION NUMBER: US 60/215, 161
: PRIOR FILING DATE: 2000-06-30
: NUMBER OF SEQ ID NOS: 8409
: SEQ ID NO 3860
: LENGTH: 83080
: TYPE: DNA
: ORGANISM: Xenorhabdus sp.
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (40567)..(40932)
: US-09-897-516-3860

Query Match 51.7%; Score 763.4; DB 6; Length 83080;
Best Local Similarity 73.4%; Pred. No. 1.9e-252;
Matches 1100; Conservative 0; Mismatches 336; Indels 43; Gaps 8;

Oy 1 gatcatggtcgaagaactaactgtgagcggtgcttaacatgcaatgcagcggtgtag 60
Db 616 gatcatggtcgaagaactaactgtgagcggtgcttaacatgcaatgcagcggtgtag 675
Oy 61 caatac-----ctagcggaacgggtgtagtaacaactgtgtaactt 102
Db 676 ggaagcaactgtgctgttctgtgtagagcggtgtagtaatgtctggtggatctg 735
Oy 103 cctcagatctgggagtaacttccgaagaaggaagttaacgggtatccgttgtagtc 162
Db 736 ccgagtggaagggtgaataacactggaacgggtgtagtaacgggtatccctcaaggga-- 793
Oy 163 acaagatttgataggttaagaattatgtctgtggaatgaacccggtgcgaattagtagt 222
Db 794 gcaaggtggggaactcttggtgcgtcaacgacatcgatgaacccagatgggttagtagt 853
Oy 223 tggtaggttaatggtctcaacaaagcgacgactgtagcggcgctgtagaggggtgtccg 282
Db 854 agtgggttaaggtgtcactctgagcgacgactcctagctgtgtgtagaggtgtagcagcc 913
Oy 283 acaatggagactggagacacgagctccatctctacgggaagcgagcaatlaagatctgtctc 342
Db 914 acaatggagactggagacacgagctccatctctacgggaagcgagcaatlaagatctgtctc 973
Oy 343 aatggggaacacccctgaagcgagcgagcgagctgaacgaagaaggtctctcgatgtgtaa 402
Db 974 aatggggaacacccctgaagcgagcgagcgagctgaacgaagaaggtctctcgatgtgtaa 1033
Oy 403 gttaactagcgagga-----aataagcagcaatgtatgtactc 447
```

Dh	1034	gtactcttcagccgggagggaagcgctgaaattgtaatacgtttccacgatctgaagcttaaccgc	1093
Oy	448	gctctaaagcaccggcttaactaactgtccagcagcccggttaatacgtatgctgcaagcgtt	507
Dh	1094	agagaagaaccaccggcttaactaccgtgcccaacacacccggttaatactcagaggtgtcaagcgtt	1153
Oy	508	gttcgcgaaatctgggcgttaaaaggtgtgcgttagccgattctgaaatctcaggtgtgtaaaact	567
Dh	1154	aattcggaaattcaacgtggcgtttaaagcgcacgcagccggtctcaataaattagatagatgaatcc	1213
Oy	568	gcgggcgtcaaaccccgctgcgcactctgaaacacaaagtcctgagatcttggtagagagcagaat	627
Dh	1214	ccgggctctaaaccctgggaactcgtacataagaactcgtttgactatagatcctctagagggggct	1273
Oy	628	gaaattccaggtgttagcgtgtgaaaatgcgttagatatactctgtagaggaacacagctgcgcgaagc	687
Dh	1274	agaattcccaagtttagccggtgaaatctcgtataagatattgttgtagagaaataccgggtgcgaagcc	1333
Oy	688	gactctgtgcctcaaaacttgacgccttgaggcagaagaacgcttggtgtgtagtaaaaggattaga	747
Dh	1334	ggcccccttgagcgaagaagacttgacgccttcaggtctgcgaagaagcgttgtagagcacaacagattaga	1393
Oy	748	tacccttgatacaccgcgcctctaaacgcttctgctac--cagttcttgagggtttcttaaccctca	806
Dh	1394	taccctgtatgctccacgcgcttaaaacgtctgaactctgtagcgttgcgtcttgagccgtgag	1453
Oy	807	gttaacggaacttaacgcgttaagtaagtgcacccgcgtgcgggagacatgctgcgcaaggtgtaaaact	866
Dh	1454	cttcctggagacttaacgcgtttaaagtctgcacccgccttgaggagtagcgtctgcgaagaattaaaacct	1513
Oy	867	aaaggaattgacagggggctccgcacaaagccggtgtagagcagatggtttcaattcgaataaacc	926
Dh	1514	aaatgtaattgagccggggccgcacaagccggtgtagagcagttggttttaattcgtatgtaaacgc	1573
Oy	927	caaaaacctcaactctgggctcttgacaatggaatctgaaatcat--gttagaataataatgagccttcg	985
Dh	1574	gaaagaaccttaacttaactcttgacaatcccaagcgaatcctctaaagatgtagggagtgccctcg	1633
Oy	986	ggcagcat--cacagcgtgcgtgcatgtttctgcgaacatcgtttctctgtagagatgttggtta	1043
Dh	1634	ggaaacgtctgagacaacgctgctgcgtcaatgtcgtctgcgtacgtctggttctgtaaatgtgttggtta	1693
Oy	1044	agtcctgcgaacgagcgcacacctcaatcgtatgctgc--taacctaaattgtagcactgct	1100
Dh	1694	agtcctgcgaacgagcgcacacctcaatcctcttctgtctgcacagcagtaatgtgtggaactcaa	1753
Oy	1101	acggaactgcgcgtgtgacaacccggtgaggaagcgggtgtagcgtccaatccctcatgacct	1160
Dh	1754	ggggagactctgcgtgtgataaacccggaagaggtgtgggtatgacgtcaagatcatcatatgacct	1813
Oy	1161	tatctccaggtcccaacacacgtctctacaaatggtccggtatagagaaggtctgcgaactctgcaaga	1220
Dh	1814	tacgagatgaggtctatacacacgtctctacaaatggtcgaagatatacaagaagagcactctgcgaga	1873
Oy	1221	ggggagctaatctcttaaaagtctggttcccagttcggattcgtgtgtcctgcgaactcgaccacatg	1280
Dh	1874	gcaagcgtgaactctataaagctctgtctgtatctccgattctggaattctgcgaactctgacatcca	1933
Oy	1281	aagctcggaatctgcgttagaataatccgcgtctcaacatgcgcgggtgtgaatagcttccggaaact	1340
Dh	1934	aagctcggaatctgcgttagaataatccgttagacaatagctcaaggtgaaatacgtctccggggcct	1993
Oy	1341	gtacacacccgcctctcacacacactatgcttggtggagacaccggagatgcttctgttaaacg	1400
Dh	1994	gtacacacccgcctctcacacacatgtagaggtgtgtgtaaaaagaagttag--gtagcttaacct	2052
Oy	1401	taaggtgacacgactcaactaaaggtgaataactgttaaaaggggttgaaatcgtctaaagaagtacc	1459
Dh	2053	tcgggtgaggtcgtctacaacttctgtatctacgtcctgggttgtagaattcgttaacaaggttaac	2111

```

Sequence 3861. Application US/09897516
GENERAL INFORMATION:
APPLICANT: Cordin, David R.
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Huesling, Joseph E.
APPLICANT: Krasomil-Osterfeld, Karina C.
APPLICANT: Malvat, Thomas M.
APPLICANT: Slater, Steven C.
APPLICANT: Spiridonov, Sergei
TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
FILE REFERENCE: 38-21(51847)B
CURRENT APPLICATION NUMBER: US/09/897,516
CURRENT FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: US 60/215, 161
PRIOR FILING DATE: 2000-06-30
NUMBER OF SEQ ID NOS: 8409
SEQ ID NO 3861
LENGTH: 83080
TYPE: DNA
ORGANISM: Xenorhabdus sp.
FEATURE:
NAME/KEY: CDS
LOCATION: (41695)..(42120)
US-09-897-516-3861

```

Query Match	51.7%	Score 763.4	DB: 6	Length 83080	
Best Local Similarity	73.4%	Pred. No. 1.9e-252			
Matches 1100	Conservative	0	Mismatches 356	Indels 43	Gaps 8
Oy	1	gatactgctcagaactaacgctgcgagcgctcttaacagatcagaagtcgagcgtag	60		
Db	616	gatactgctcagaatltgaaacgctgcgagcgctcaacacatgcgaagtcgagcgtag	675		
Oy	61	caatac-----ctagcgcgaaacgggtgaagtaacgtygtactctt	102		
Db	676	ggaagcaagcttgctgctcttgctgaaacgagcgcgagcggtgtgaatctgctggtgacttg	735		
Oy	103	ctctcgagcttgaggtaactcttcgaaaggaagtaacgggtagtctctgttgatc	162		
Db	736	ccgagctgaggggtgataacacacactgtgaaacgctgcttaataacggcaltgacatcttaag	793		
Oy	163	acaagaatttgatagataaagatttcttgcttggaatltagcccgcgccgagataagtagt	222		
Db	794	gcaaggtgagggaacttcgggcttaacgcgcacatcgagatgaacccagatggtatagtagt	853		
Oy	223	tgtgtgagtlaatgcttcaaccaaagcgagatctgttagccgcttgagaggggtgtccggtc	282		
Db	854	aggtggtgtaaacgctcacctcactagcgagatccctagctgtgcttgagagatgacacagcc	913		
Oy	283	acaatgnaactgagagcaagctgcatctactctctcgggagggacaaatltgaatctgtctc	342		
Db	914	acaatcggaacttgagacaaagcccaagatctctcagggagagccaaatgtgggaatatltgcac	973		
Oy	343	aatggtggaaaccccgaaagcaagcgacgcgcgttgaaaggaagaaggtctctcgatgtgtaa	402		
Db	974	aatgtggtcgaaagcccgatgagcagcatctgcgcgtgtcatgaaaggaagccttcgggtgttaa	1033		
Oy	403	gttcatltaggcaagaaa-----aataagcagcaatgtgatatgtatcct	447		
Db	1034	gtacttcttcagcgggagagagcgctgaaagtltgaatacgttctcaagaltgacatccccc	1093		
Oy	448	gcttaagaugacccggtaactacgttgcacgaacacggcgtaatacgtatgttcaagcgtt	507		
Db	1094	agaagaagaacccggtaactccgttcgcacgaacggcggtgttaatacggagaggtgtcaagcgtt	1153		
Oy	508	gttcggaatcatcttggtgcgttaaaagggtgcgttagggcggaattgtgaatcaggtgtgaaact	567		
Db	1154	aatcggaattacttggtgcgttaaaagcgacgcagcgatggttcaattgaattgagtgaatccc	1213		
Oy	568	gcgggctcaaccgttgctgtcacttgaaactacaagcttgagatttggagaggcaagt	627		

Db 1214 ccgggcttaacctggaactgcatacgaactcgttgactagatctcgttagagggggt 1273  
Qy 628 ggaattccaggtgtagcgggtgaatctgtaatactcgtgagaacaccagtcggcgaagc 687  
Db 1274 agaatccacgtgtagcgttgaaatgcgttagagaatgtgaggaataaccgttgcgaagc 1333  
Qy 688 gacttgcgtgctcaaaactgcgttcgaagcgtggaagcgttggttagtaaacgggattaga 747  
Db 1334 gggccctcgagcgaagactgcgttcgaagcgttgggaagcgtggaagcgaagattaga 1393  
Qy 748 taaccggttaatccacgacctaaacgttctac-cagtgtctgggggtttaacctca 806  
Db 1394 taacctgttagtccacgacctgaacgactgtcgaactgttggtcttgagccgtgg 1453  
Qy 807 gtaacgaacctaaacgttgaatgacccgttggggaactatgcttcgcaaggtgaactc 866  
Db 1454 ctccggagctaaacgcttaagctcgaacctcggggaactcgttcgcaagattaaactc 1513  
Qy 867 aaaggaattgacgggggtcgcacaaagcgtggaagcgttggttaattcgaataacc 926  
Db 1514 aaatgaattgacgggggtcgcacaaagcgtggaagcgttggttaattcgaataacc 1573  
Qy 927 caaaaacctcaacctggtgacatgacatcgaatcat-gtagagataatgaacctcg 985  
Db 1574 gaagaaccttaaccttgaacctgaacctgaacctcgaagatgaaggaatgccttcg 1633  
Qy 986 ggcagatt--cacagggtgctgacgtgtgctgaagcgtgctgctgagatgttggtta 1043  
Db 1634 ggaacgctgagacaggtgtcgtcgaagcgtgctgctgagatgttggtta 1693  
Qy 1044 agtcccgcaagcgaagcgaacctatcgtcatgttcg---taacctaaatgtggcaactgt 1100  
Db 1694 agtcccgcaagcgaagcgaacctatccttctgttcgcaagcgtatgtgtggaaactca 1753  
Qy 1101 acgaactgcgtgcgaacacccggaaggaagcgtggaatgacgtcaaatcctcatgacct 1160  
Db 1754 gggagactgcgtgtgataaacccggaaggaagcgtggaatgacgtcaaatcctcatgacct 1813  
Qy 1161 tatgccaggggcacaacacgtgtcgaataatggtccgaataagaggtcggcaactcgcgaaga 1220  
Db 1814 taagagtaagggtcacaacacgtgtcgaataatggtccgaataagaggtcggcaactcgcgaaga 1873  
Qy 1221 gggagctaatctctaaagctcgttccagctcgtggttggttgcgaactcgaacctca 1280  
Db 1874 gcaagcgaactcatalaagctgtctgtagtccggaatggaatgcgaactcgaacctca 1933  
Qy 1281 aagtcggatcgtcgtatcgtcgtggaatcgaatgcgtggaatgcgttcccggaactc 1340  
Db 1934 aagtcggatcgtcgtatcgtcgtggaatcgaatgcgttcccggaactc 1993  
Qy 1341 gtaacacccgctcacaacacccggaatggtgggaacccggaatggtcttgaacctg 1400  
Db 1994 gtaacacccgctcacaacacccggaatggtgggaacccggaatggtcttgaacctg 2052  
Qy 1401 taaggagacagactactaagggtgaactcgttaaggggtgaagtcgttaacaggtacc 1459  
Db 2053 tcgggagggcgttaaccttgtgattcatgacgtgggtgaagtcgttaacaggttaac 2111

Search completed: December 15, 2001, 05:58:16  
Job time: 12259 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 05:42:50 : Search time 7962.77 Seconds  
(without alignments)  
48.423 Million cell updates/sec

Title: US-09-380-826a-2

Perfect score: 22

Sequence: 1 tfttgatcacacagattgata 22

Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 17159718 seqs, 8763200856 residues

Total number of hits satisfying chosen parameters: 34319436

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_NA\_Main:\*

1: /cgn2\_6/ptodata/2/pna/US07\_PCTUS\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/pna/US08\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/pna/US07\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/pna/US081\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/pna/US081\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/pna/US083\_COMB.seq:\*  
7: /cgn2\_6/ptodata/2/pna/US083\_COMB.seq:\*  
8: /cgn2\_6/ptodata/2/pna/US084\_COMB.seq:\*  
9: /cgn2\_6/ptodata/2/pna/US085\_COMB.seq:\*  
10: /cgn2\_6/ptodata/2/pna/US086\_COMB.seq:\*  
11: /cgn2\_6/ptodata/2/pna/US087\_COMB.seq:\*  
12: /cgn2\_6/ptodata/2/pna/US088\_COMB.seq:\*  
13: /cgn2\_6/ptodata/2/pna/US089\_COMB.seq:\*  
14: /cgn2\_6/ptodata/2/pna/US090\_COMB.seq:\*  
15: /cgn2\_6/ptodata/2/pna/US091\_COMB.seq:\*  
16: /cgn2\_6/ptodata/2/pna/US093\_COMB.seq:\*  
17: /cgn2\_6/ptodata/2/pna/US093\_COMB.seq:\*  
18: /cgn2\_6/ptodata/2/pna/US094\_COMB.seq:\*  
19: /cgn2\_6/ptodata/2/pna/US095\_COMB.seq:\*  
20: /cgn2\_6/ptodata/2/pna/US095\_COMB.seq:\*  
21: /cgn2\_6/ptodata/2/pna/US095\_COMB.seq:\*  
22: /cgn2\_6/ptodata/2/pna/US095\_COMB.seq:\*  
23: /cgn2\_6/ptodata/2/pna/US096\_COMB.seq:\*  
24: /cgn2\_6/ptodata/2/pna/US096\_COMB.seq:\*  
25: /cgn2\_6/ptodata/2/pna/US096\_COMB.seq:\*  
26: /cgn2\_6/ptodata/2/pna/US096\_COMB.seq:\*  
27: /cgn2\_6/ptodata/2/pna/US096\_COMB.seq:\*  
28: /cgn2\_6/ptodata/2/pna/US096\_COMB.seq:\*  
29: /cgn2\_6/ptodata/2/pna/US097\_COMB.seq:\*  
30: /cgn2\_6/ptodata/2/pna/US097\_COMB.seq:\*  
31: /cgn2\_6/ptodata/2/pna/US098\_COMB.seq:\*  
32: /cgn2\_6/ptodata/2/pna/US098\_COMB.seq:\*  
33: /cgn2\_6/ptodata/2/pna/US098\_COMB.seq:\*  
34: /cgn2\_6/ptodata/2/pna/US098\_COMB.seq:\*  
35: /cgn2\_6/ptodata/2/pna/US098\_COMB.seq:\*  
36: /cgn2\_6/ptodata/2/pna/US098\_COMB.seq:\*  
37: /cgn2\_6/ptodata/2/pna/US098\_COMB.seq:\*  
38: /cgn2\_6/ptodata/2/pna/US098\_COMB.seq:\*  
39: /cgn2\_6/ptodata/2/pna/US098\_COMB.seq:\*  
40: /cgn2\_6/ptodata/2/pna/US098\_COMB.seq:\*  
41: /cgn2\_6/ptodata/2/pna/US098\_COMB.seq:\*  
42: /cgn2\_6/ptodata/2/pna/US098\_COMB.seq:\*  
43: /cgn2\_6/ptodata/2/pna/US098\_COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Match	Query length	ID	Description
1	22	100.0	22	US-09-380-826a-2	Sequence 2, Appl
2	22	100.0	17	US-09-380-826a-7	Sequence 1, Appl
3	22	100.0	17	US-09-380-826a-11944	Sequence 1, Appl
4	18.8	85.5	420	US-09-649-1654-7425	Sequence 1944, A
5	18.8	85.5	581	US-09-654-617-175385	Sequence 175385, A
6	18.8	85.5	581	US-09-654-617-175385	Sequence 175385, A
7	18.8	85.5	581	US-09-654-617-175385	Sequence 175385, A
8	18.8	85.5	649	US-09-565-309A-20908	Sequence 20908, A
9	18.8	85.5	1069	US-09-565-309A-57118	Sequence 57118, A
10	18.8	85.5	1072	US-09-565-309A-49885	Sequence 49885, A
11	18.8	85.5	1128	US-09-654-617-17747	Sequence 117747, A
12	18.8	85.5	1128	US-09-654-617-17747	Sequence 117747, A
13	18.8	85.5	1429	US-09-138-103-11564	Sequence 11564, A
14	18.8	85.5	5874	US-09-404-520-5289	Sequence 5289, A
15	18.8	85.5	92624	US-09-534-859-352	Sequence 352, A
16	18.8	85.5	92624	US-09-803-736-352	Sequence 352, A
17	18.4	83.6	2101	US-09-143-993-148	Sequence 148, A
18	18.4	83.6	2463	US-09-191-637-27704	Sequence 27704, A
19	18.4	83.6	2463	US-09-191-637-27704	Sequence 27704, A
20	18.4	83.6	2486	US-09-167-217-19072	Sequence 22313, A
21	18.4	83.6	2486	US-09-167-217-19072	Sequence 18072, A
22	18.4	83.6	3442	US-09-191-637-19008	Sequence 19008, A
23	18.4	83.6	3831	US-09-167-217-18350	Sequence 18350, A
24	18.4	83.6	4637	US-09-619-049-937	Sequence 937, A
25	18.4	83.6	4637	US-09-171-637-15063	Sequence 1465, A
26	18.4	83.6	4637	US-09-173-464-15063	Sequence 15063, A
27	18.4	83.6	4637	US-09-191-637-18351	Sequence 18351, A
28	18.4	83.6	4637	US-09-191-637-18351	Sequence 18351, A
29	18.4	83.6	4637	US-09-191-637-18351	Sequence 18351, A
30	18.4	83.6	4637	US-09-191-637-18351	Sequence 18351, A
31	18.4	83.6	4637	US-09-191-637-18351	Sequence 18351, A
32	18.4	83.6	4637	US-09-191-637-18351	Sequence 18351, A
33	18.4	83.6	4637	US-09-191-637-18351	Sequence 18351, A
34	18.4	83.6	4637	US-09-191-637-18351	Sequence 18351, A
35	18.4	83.6	4637	US-09-191-637-18351	Sequence 18351, A
36	18.4	83.6	4637	US-09-191-637-18351	Sequence 18351, A
37	18.4	83.6	4637	US-09-191-637-18351	Sequence 18351, A
38	18.4	83.6	4637	US-09-191-637-18351	Sequence 18351, A
39	18.4	83.6	4637	US-09-191-637-18351	Sequence 18351, A
40	18.4	83.6	4637	US-09-191-637-18351	Sequence 18351, A
41	18.4	83.6	4637	US-09-191-637-18351	Sequence 18351, A
42	18.4	83.6	4637	US-09-191-637-18351	Sequence 18351, A
43	18.4	83.6	4637	US-09-191-637-18351	Sequence 18351, A

42 17.2 78.2 249 24 US-09-638-258-520 Sequence 520, App  
43 17.2 78.2 249 32 US-09-924-035-520 Sequence 520, App  
44 17.2 78.2 278 25 US-09-654-617-370318 Sequence 370318,  
45 17.2 78.2 278 27 US-09-684-016-370318 Sequence 370318,

## ALIGNMENTS

RESULT 1  
US-09-380-826A-2  
; Sequence 2, Application US/09380826A  
; GENERAL INFORMATION:  
; APPLICANT: Chappell, Rod  
; TITLE OF INVENTION: LEPTOSPIRA PATHOGENS  
; FILE REFERENCE: DAVIE79.001APC  
; CURRENT APPLICATION NUMBER: US/09/380,826A  
; CURRENT FILING DATE: 1999-11-22  
; PRIOR APPLICATION NUMBER: PCT/AU98/00145  
; PRIOR FILING DATE: 1998-03-06  
; PRIOR APPLICATION NUMBER: AU P05494/97  
; PRIOR FILING DATE: 1997-03-07  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 22  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Primer for L. fainei.  
US-09-380-826A-2

Query Match 100.0%; Score 22; DB 17; Length 22;  
Best Local Similarity 100.0%; Pred. No. 2;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 tgttgatcacacaagatttgata 22  
|||  
Db 1 tgttgatcacacaagatttgata 22  
RESULT 2  
US-09-380-826A-7  
; Sequence 7, Application US/09380826A  
; GENERAL INFORMATION:  
; APPLICANT: Chappell, Rod  
; TITLE OF INVENTION: LEPTOSPIRA PATHOGENS  
; FILE REFERENCE: DAVIE79.001APC  
; CURRENT APPLICATION NUMBER: US/09/380,826A  
; CURRENT FILING DATE: 1999-11-22  
; PRIOR APPLICATION NUMBER: PCT/AU98/00145  
; PRIOR FILING DATE: 1998-03-06  
; PRIOR APPLICATION NUMBER: AU P05494/97  
; PRIOR FILING DATE: 1997-03-07  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 22  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Primer for L. fainei.  
US-09-380-826A-7

Query Match 100.0%; Score 22; DB 17; Length 22;  
Best Local Similarity 100.0%; Pred. No. 2;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 tgttgatcacacaagatttgata 22  
|||  
Db 1 tgttgatcacacaagatttgata 22

RESULT 3  
US-09-380-826A-1  
; Sequence 1, Application US/09380826A  
; GENERAL INFORMATION:  
; APPLICANT: Chappell, Rod  
; TITLE OF INVENTION: LEPTOSPIRA PATHOGENS  
; FILE REFERENCE: DAVIE79.001APC  
; CURRENT APPLICATION NUMBER: US/09/380,826A  
; CURRENT FILING DATE: 1999-11-22  
; PRIOR APPLICATION NUMBER: PCT/AU98/00145  
; PRIOR FILING DATE: 1998-03-06  
; PRIOR APPLICATION NUMBER: AU P05494/97  
; PRIOR FILING DATE: 1997-03-07  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 1477  
; TYPE: DNA  
; ORGANISM: Leptospira fainei  
US-09-380-826A-1

Query Match 100.0%; Score 22; DB 17; Length 1477;  
Best Local Similarity 100.0%; Pred. No. 3.3;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 tgttgatcacacaagatttgata 22  
|||  
Db 154 tgttgatcacacaagatttgata 175

RESULT 4  
US-60-132-861-11944  
; Sequence 11944, Application US/60132861  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: ASPERGILLUS NIDULANS GENOME SEQUENCE AND USES THEREOF  
; FILE REFERENCE: 38-21(115485)C  
; CURRENT APPLICATION NUMBER: US/60/132,861  
; CURRENT FILING DATE: 1999-05-07  
; NUMBER OF SEQ ID NOS: 13800  
; SEQ ID NO 11944  
; LENGTH: 420  
; TYPE: DNA  
; ORGANISM: Aspergillus nidulans  
US-60-132-861-11944

Query Match 85.5%; Score 18.8; DB 46; Length 420;  
Best Local Similarity 90.9%; Pred. No. 1.1e+02;  
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
OY 1 tgttgatcacacaagatttgata 22  
|||  
Db 234 tgttgatcacacaagatttgata 255

RESULT 5  
US-09-649-165A-7425/C  
; Sequence 7425, Application US/09649165A  
; GENERAL INFORMATION:  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Doyle, Martin  
; APPLICANT: Momiyama, Monika  
; APPLICANT: Wang, Xinhao  
; TITLE OF INVENTION: DETECTING EXPRESSION OF ARABIDOPSIS THALIANA GENES  
; FILE REFERENCE: PA-0019 US  
; CURRENT APPLICATION NUMBER: US/09/649,165A  
; CURRENT FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: 60/150,517



PRIOR FILING DATE: 1999-08-24  
NUMBER OF SEQ ID NOS: 7753  
SOFTWARE: PERL Program  
SEQ ID NO 7425  
LENGTH: 530  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
FEATURE:  
NAME/KEY: misc.feature  
OTHER INFORMATION: Incyte ID No: 701676626  
LOCATION: 6  
OTHER INFORMATION: a, t, c, g, or other  
US-09-649-165A-7425

Query Match 85.5%; Score 18.8; DB 25; Length 530;  
Best Local Similarity 90.9%; Pred. No. 1.1e+02;  
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 tgttgatcacagaattgata 22  
||||| |||||||||  
DB 254 tgttgatcacagaattgaca 233

RESULT 6  
US-09-654-617-175385  
Sequence 175385, Application US/09654617  
GENERAL INFORMATION:  
APPLICANT: Kovacic, David K.  
APPLICANT: Liu, Jingdong  
TITLE OF INVENTION: Annotated Plant Genes  
FILE REFERENCE: 38-21(15097)D  
CURRENT APPLICATION NUMBER: US/09/654,617  
CURRENT FILING DATE: 2000-09-05  
NUMBER OF SEQ ID NOS: 463173  
SEQ ID NO 175385  
LENGTH: 581  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
US-09-654-617-175385

Query Match 85.5%; Score 18.8; DB 25; Length 581;  
Best Local Similarity 90.9%; Pred. No. 1.1e+02;  
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 tgttgatcacagaattgata 22  
||||| |||||||||  
DB 420 tgttgatcacagaattgaca 441

RESULT 7  
US-09-684-016-175385  
Sequence 175385, Application US/09684016  
GENERAL INFORMATION:  
APPLICANT: Liu, Jingdong  
APPLICANT: Kovacic, David K.  
TITLE OF INVENTION: Annotated Plant Genes  
FILE REFERENCE: 38-21(15097)D  
CURRENT APPLICATION NUMBER: US/09/684,016  
CURRENT FILING DATE: 2000-10-10  
PRIOR APPLICATION NUMBER: US 09/654,617  
PRIOR FILING DATE: 2000-09-05  
NUMBER OF SEQ ID NOS: 463173  
SEQ ID NO 175385  
LENGTH: 581  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
US-09-684-016-175385

Query Match 85.5%; Score 18.8; DB 27; Length 581;

Best Local Similarity 90.9%; Pred. No. 1.1e+02;  
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 tgttgatcacagaattgata 22  
||||| |||||||||  
DB 420 tgttgatcacagaattgaca 441

RESULT 8  
US-09-565-309A-20908  
Sequence 20908, Application US/09565309A  
GENERAL INFORMATION:  
APPLICANT: ALEXANDROV, Nikolai  
APPLICANT: BROVER, Vyacheslav  
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
FILE REFERENCE: 2750-0853P  
CURRENT APPLICATION NUMBER: US/09/565,309A  
CURRENT FILING DATE: 2000-05-05  
NUMBER OF SEQ ID NOS: 68449  
SEQ ID NO 20908  
LENGTH: 649  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: (1)..(649)  
OTHER INFORMATION: any n = a, g, c, t, unknown, or other  
NAME/KEY: misc.feature  
LOCATION: (1)..(649)  
OTHER INFORMATION: 27496:110925 (Clone Number:Unique Sequence Identifier)  
US-09-565-309A-20908

Query Match 85.5%; Score 18.8; DB 22; Length 649;  
Best Local Similarity 90.9%; Pred. No. 1.1e+02;  
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 tgttgatcacagaattgata 22  
||||| |||||||||  
DB 265 tgttgatcacagaattgaca 286

RESULT 9  
US-09-565-309A-57118  
Sequence 57118, Application US/09565309A  
GENERAL INFORMATION:  
APPLICANT: ALEXANDROV, Nikolai  
APPLICANT: BROVER, Vyacheslav  
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID  
FILE REFERENCE: 2750-0853P  
CURRENT APPLICATION NUMBER: US/09/565,309A  
CURRENT FILING DATE: 2000-05-05  
NUMBER OF SEQ ID NOS: 68449  
SEQ ID NO 57118  
LENGTH: 1069  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: (1)..(1069)  
OTHER INFORMATION: any n = a, g, c, t, unknown, or other  
NAME/KEY: misc.feature  
LOCATION: (1)..(1069)  
OTHER INFORMATION: 27496 : OVERLAP (Clone Number : OVERLAP)  
US-09-565-309A-57118

Query Match 85.5%; Score 18.8; DB 22; Length 1069;  
Best Local Similarity 90.9%; Pred. No. 1.2e+02;  
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tttggtacacaagatttgata 22  
||||| |||||||||  
Db 690 tttggtacacaagatttgaca 711

## RESULT 10

US-09-565-309A-49885  
; Sequence 49885, Application US/09565309A  
; GENERAL INFORMATION:  
; APPLICANT: ALEXANDROV, Nickolai  
; APPLICANT: BROVER, Vyacheslav  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; TITLE OF INVENTION: THEREBY  
; FILE REFERENCE: 2750-0853P  
; CURRENT APPLICATION NUMBER: US/09/565,309A  
; CURRENT FILING DATE: 2000-05-05  
; NUMBER OF SEQ ID NOS: 68449  
; SEQ ID NO 49885  
; LENGTH: 1072  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc-feature  
; LOCATION: (1)..(1072)  
; OTHER INFORMATION: any n = a, g, c, t, unknown, or other  
; NAME/KEY: misc-feature  
; LOCATION: (1)..(1072)  
; OTHER INFORMATION: 27496 : 5TAG CONSENSUS (Clone Number:5tag\_consensus)  
US-09-565-309A-49885

Query Match 85.5%; Score 18.8; DB 22; Length 1072;  
Best Local Similarity 90.9%; Pred. No. 1.2e+02;  
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
OY 1 tttggtacacaagatttgata 22  
||||| |||||||||  
Db 691 tttggtacacaagatttgaca 712

## RESULT 11

US-09-654-617-117747  
; Sequence 117747, Application US/09654617  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: LIU, Jindong  
; TITLE OF INVENTION: Annotated Plant Genes  
; FILE REFERENCE: 38-21(15097)D  
; CURRENT APPLICATION NUMBER: US/09/654,617  
; CURRENT FILING DATE: 2000-09-05  
; NUMBER OF SEQ ID NOS: 463173  
; SEQ ID NO 117747  
; LENGTH: 1128  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
; OTHER INFORMATION: unsure at all n locations  
US-09-654-617-117747

Query Match 85.5%; Score 18.8; DB 25; Length 1128;  
Best Local Similarity 90.9%; Pred. No. 1.2e+02;  
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
OY 1 tttggtacacaagatttgata 22  
||||| |||||||||  
Db 719 tttggtacacaagatttgaca 740

RESULT 12  
US-09-684-016-117747  
; Sequence 117747, Application US/09684016  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.

APPLICANT: LIU, Jindong Annotated Plant Genes  
; TITLE OF INVENTION: 38-21(15097)D  
; FILE REFERENCE: 38-21(15097)D  
; CURRENT APPLICATION NUMBER: US/09/684,016  
; CURRENT FILING DATE: 2000-10-10  
; PRIOR APPLICATION NUMBER: US 09/654,617  
; PRIOR FILING DATE: 2000-09-05  
; NUMBER OF SEQ ID NOS: 463173  
; SEQ ID NO 117747  
; LENGTH: 1128  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(1128)  
; OTHER INFORMATION: unsure at all n locations  
US-09-684-016-117747

Query Match 85.5%; Score 18.8; DB 27; Length 1128;  
Best Local Similarity 90.9%; Pred. No. 1.2e+02;  
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
OY 1 tttggtacacaagatttgata 22  
||||| |||||||||  
Db 719 tttggtacacaagatttgaca 740

## RESULT 13

US-60-138-103-11564/C  
; Sequence 11564, Application US/60138103  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Timberlake, William E.  
; TITLE OF INVENTION: ASPERGILLUS NIDULANS GENOME SEQUENCE AND USES THEREOF  
; FILE REFERENCE: 38-10(15485)D  
; CURRENT APPLICATION NUMBER: US/60/138,103  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 28006  
; SEQ ID NO 11564  
; LENGTH: 1429  
; TYPE: DNA  
; ORGANISM: Aspergillus nidulans  
US-60-138-103-11564

Query Match 85.5%; Score 18.8; DB 46; Length 1429;  
Best Local Similarity 90.9%; Pred. No. 1.2e+02;  
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
OY 1 tttggtacacaagatttgata 22  
||||| |||||||||  
Db 1206 tttggtacacaagatttgaca 1185

## RESULT 14

US-09-404-520-5289  
; Sequence 5289, Application US/09404520  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Chodas, Azita  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: McIninch, James  
; APPLICANT: Timberlake, William E.  
; APPLICANT: Yu, Jaehyuk  
; TITLE OF INVENTION: EMERICELLA NIDULANS GENOME SEQUENCE AND USES THEREOF  
; FILE REFERENCE: 38-10(15498)A  
; CURRENT APPLICATION NUMBER: US/09/404,520  
; CURRENT FILING DATE: 1999-09-23  
; NUMBER OF SEQ ID NOS: 44345  
; SEQ ID NO 5289  
; LENGTH: 5874  
; TYPE: DNA

ORGANISM: Aspergillus nidulans  
US-09-404-520-5289

Query Match 85.5%; Score 18.8; DB 18; Length 5874;  
Best Local Similarity 90.9%; Pred. No. 1.4e+02;  
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tgttgatcacaagatttgata 22  
||||| ||||||| ||||  
DB 248 tgttgatcacaagatttgata 269

RESULT 15

US-09-534-859-352  
Sequence 352, Application US/09534859  
GENERAL INFORMATION:  
APPLICANT: Bush, David F.  
APPLICANT: Last, Robert L.  
APPLICANT: Levin, Irena M.  
APPLICANT: Norris, Susan R.  
APPLICANT: Parnell, Laurence D.  
APPLICANT: Rounsley, Steven D.  
APPLICANT: Wiegand, Roger C.  
TITLE OF INVENTION: PLANT POLYMORPHIC MARKERS AND USES THEREOF  
FILE REFERENCE: 38-10(15493)B  
CURRENT APPLICATION NUMBER: US/09/534,859  
CURRENT FILING DATE: 2000-03-29  
NUMBER OF SEQ ID NOS: 1127  
SEQ ID NO 352  
LENGTH: 92624  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
US-09-534-859-352

Query Match 85.5%; Score 18.8; DB 20; Length 92624;  
Best Local Similarity 90.9%; Pred. No. 2e+02;  
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tgttgatcacaagatttgata 22  
||||| ||||||| |||||  
DB 35807 tgttgatcacaagatttgata 35828

Search completed: December 15, 2001, 05:42:55  
Job time: 15820 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 05:58:16 : Search time 474.83 Seconds  
(without alignments)  
66.629 Million cell updates/sec

Title: US-09-380-826A-2

Perfect score: 22

Sequence: 1 tttgtgacacaaagattgata 22

Scoring table: IDENTITY\_NUC

Searched: Gapop 10.0, Gapext 1.0

Total number of hits satisfying chosen parameters: 1946350

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 0%

Listing first 45 summaries

Database : Pending\_Patents\_NA\_New:\*  
1: /cgn2\_6/ptodata/1/pna/PCF\_NEW\_COMB.seq:\*  
2: /cgn2\_6/ptodata/1/pna/US06\_NEW\_COMB.seq:\*  
3: /cgn2\_6/ptodata/1/pna/US07\_NEW\_COMB.seq:\*  
4: /cgn2\_6/ptodata/1/pna/US08\_NEW\_COMB.seq:\*  
5: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq:\*  
6: /cgn2\_6/ptodata/1/pna/US09\_NEW\_COMB.seq1:\*  
7: /cgn2\_6/ptodata/1/pna/US10\_NEW\_COMB.seq:\*  
8: /cgn2\_6/ptodata/1/pna/US60\_NEW\_COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18.8	85.5	1069	6 US-09-620-394B-4668	Sequence 4668, Ap
2	17.2	78.2	249	5 US-09-924-035A-520	Sequence 520, App
3	17.2	78.2	622	6 US-09-985-678-211824	Sequence 211824, A
4	17.2	78.2	2038	6 US-09-922-279A-2773	Sequence 2773, Ap
5	17.2	78.2	46891	6 US-09-815-264-82805	Sequence 82805, A
6	16.8	76.4	267	6 US-09-982-402-1036	Sequence 1036, Ap
7	16.8	76.4	551	6 US-09-849-526A-27340	Sequence 27340, A
8	16.8	76.4	849	6 US-09-815-264-26466	Sequence 26466, A
9	16.8	76.4	1246	6 US-09-815-264-91672	Sequence 91672, A
10	16.8	76.4	1501	6 US-09-815-264-23140	Sequence 23140, A
11	16.8	76.4	1501	6 US-09-815-264-46882	Sequence 46882, A
12	16.8	76.4	1501	6 US-09-815-264-53250	Sequence 53250, A
13	16.8	76.4	1501	6 US-09-815-264-71944	Sequence 71944, A
14	16.8	76.4	12081	6 US-09-815-264-76889	Sequence 76889, A
15	16.8	76.4	12198	6 US-09-815-264-76553	Sequence 76553, A
16	16.8	76.4	41269	6 US-09-815-264-59808	Sequence 59808, A
17	16.8	76.4	50929	6 US-09-815-264-61033	Sequence 61033, A
18	16.8	76.4	67216	6 US-09-815-264-71536	Sequence 71536, A
19	16.8	76.4	360	6 US-09-985-678-279880	Sequence 279880, A
20	16.4	74.5	4861	6 US-09-815-264-109466	Sequence 109466, A
21	16.4	74.5	278	6 US-09-985-678-279880	Sequence 279880, A
22	16.2	73.6	376	6 US-09-912-293-18469	Sequence 18469, A
23	16.2	73.6	377	6 US-09-845-487A-238	Sequence 238, App
24	16.2	73.6	419	6 US-09-933-524A-2094	Sequence 2094, Ap
25	16.2	73.6			

## ALIGNMENTS

26	16.2	73.6	461	6 US-09-849-526A-12135	Sequence 12135, A
27	16.2	73.6	845	6 US-09-922-279A-1268	Sequence 1268, Ap
28	16.2	73.6	1251	6 US-09-898-888A-9009	Sequence 9009, Ap
29	16.2	73.6	1457	6 US-09-815-264-60838	Sequence 60838, A
30	16.2	73.6	1501	6 US-09-815-264-39775	Sequence 39775, A
31	16.2	73.6	1501	6 US-09-815-264-39776	Sequence 39776, A
32	16.2	73.6	1501	6 US-09-815-264-43531	Sequence 43531, A
33	16.2	73.6	3603	6 US-09-815-264-80803	Sequence 80803, A
34	16.2	73.6	6237	6 US-09-815-264-72243	Sequence 72243, A
35	16.2	73.6	7856	6 US-09-815-264-86401	Sequence 86401, A
36	16.2	73.6	9723	6 US-09-815-264-75693	Sequence 75693, A
37	16.2	73.6	10839	6 US-09-815-264-78946	Sequence 78946, A
38	16.2	73.6	26601	6 US-09-815-264-74669	Sequence 74669, A
39	16.2	73.6	32005	6 US-09-815-264-73015	Sequence 73015, A
40	16.2	73.6	36227	6 US-09-815-264-66819	Sequence 66819, A
41	16	72.7	610	6 US-09-985-678-213635	Sequence 213635, A
42	16	72.7	1381	6 US-09-815-264-108378	Sequence 108378, A
43	16	72.7	37658	6 US-09-815-264-74259	Sequence 74259, A
44	16	72.7	39365	6 US-09-815-264-79204	Sequence 79204, A
45	15.8	71.8	270	6 US-09-985-678-42368	Sequence 42368, A

```

RESULT 1
US-09-620-394B-4668
; Sequence 4668, Application US/09620394B
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai
; TITLE OF INVENTION: Brover, Vyacheslav
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptid
; FILE REFERENCE: 2750-1067P
; CURRENT APPLICATION NUMBER: US/09/620,394B
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 9131
; SEQ ID NO 4668
; LENGTH: 1069
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..1069
; OTHER INFORMATION: any n = a, g, c, t, unknown, or other
; NAME/KEY: misc.feature
; LOCATION: 1..1069
; OTHER INFORMATION: Ceres Seq. ID 1393173
US-09-620-394B-4668

Query Match 85.5%; Score 18.8; DB 6; Length 1069;
Best Local Similarity 90.9%; Pred. No. 16;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 690 tttgtgacacaaagattgata 22
||||| |||||||||
1 tttgtgacacaaagattgata 711

RESULT 2
US-09-924-035A-520
; Sequence 520, Application US/09924035A
; GENERAL INFORMATION:
; APPLICANT: Grilach, Jrn
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; TITLE OF INVENTION: thaliana
; FILE REFERENCE: 2011US
; CURRENT APPLICATION NUMBER: US/09/924,035A
; CURRENT FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: US 60/148,784
; PRIOR FILING DATE: 1999-08-13
; NUMBER OF SEQ ID NOS: 900

```

```

; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 520
; LENGTH: 249
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(249)
; OTHER INFORMATION: n = A,T,C or G
US-09-924-035A-520

Query Match          78.2%; Score 17.2; DB 5; Length 249;
Best Local Similarity 86.4%; Pred. No. 81;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 tgttgatcacaaagattgata 22
||||| | ||||| ||||| ||
Db 54 tgttgatcacaaagattgata 75

RESULT 3
US-09-985-678-211824
; Sequence 211824, Application US/09985678
; GENERAL INFORMATION:
; APPLICANT: Cheikh, Norline
; APPLICANT: Liu, Jingdong
; TITLE OF INVENTION: Annotated Plant Genes
; FILE REFERENCE: 16517.255/38-21(15097)F
; CURRENT APPLICATION NUMBER: US/09/985,678
; CURRENT FILING DATE: 2001-11-05
; PRIOR APPLICATION NUMBER: US 09/304,517
; PRIOR FILING DATE: 1999-05-06
; NUMBER OF SEQ ID NOS: 295529
; SEQ ID NO 211824
; LENGTH: 622
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)...(622)
; OTHER INFORMATION: unsure at all n locations
US-09-985-678-211824

Query Match          78.2%; Score 17.2; DB 6; Length 622;
Best Local Similarity 86.4%; Pred. No. 87;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 tgttgatcacaaagattgata 22
||||| | ||||| ||||| ||
Db 520 tgttgatcacaaagattgata 541

RESULT 4
US-09-922-279A-2773
; Sequence 2773, Application US/09922279A
; GENERAL INFORMATION:
; APPLICANT: Tang, Yuanhua T.
; APPLICANT: Tillinghast, John
; APPLICANT: Slinku, Ankura
; APPLICANT: Liu, Chenghua
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Contigs Obtained
; TITLE OF INVENTION: From Various Libraries
; FILE REFERENCE: 785
; CURRENT APPLICATION NUMBER: US/09/922,279A
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/491,404
; PRIOR FILING DATE: 2000-01-25
; NUMBER OF SEQ ID NOS: 3796
; SOFTWARE: PF_SP_genes Version 1.0
; SEQ ID NO 2773

```

```

; LENGTH: 2038
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: (1091)...(956)
; OTHER INFORMATION: this location contains the signal peptide sequence,
; OTHER INFORMATION: MYLFLPFGVYSIRSLFLGRPIRGITSDFTLIFSNLDSMPLS, Run with Signal
; NAME/KEY: misc_feature
; LOCATION: (605)...(1902)
; OTHER INFORMATION: similar to g15931821 in the genepept database release 114,
; OTHER INFORMATION: Run with FASTXY 3.3f00, default parameters
US-09-922-279A-2773

Query Match          78.2%; Score 17.2; DB 6; Length 2038;
Best Local Similarity 86.4%; Pred. No. 96;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 tgttgatcacaaagattgata 22
||||| | ||||| ||||| ||
Db 1929 tgttgatcacaaagattgata 1950

RESULT 5
US-09-815-264-82805
; Sequence 82805, Application US/09815264
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Cao, Yongwei
; APPLICANT: Dotson, Stanton B.
; APPLICANT: Koshi, Jeffrey M.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: McIninch, James
; APPLICANT: Wu, Mei
; TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51237)G
; CURRENT APPLICATION NUMBER: US/09/815,264
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 09/620,392
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/702,134
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 109669
; SEQ ID NO 82805
; LENGTH: 46891
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-815-264-82805

Query Match          78.2%; Score 17.2; DB 6; Length 46891;
Best Local Similarity 86.4%; Pred. No. 1,3e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 tgttgatcacaaagattgata 22
||||| | ||||| ||||| ||
Db 45218 tgttgatcacaaagattgata 45239

RESULT 6
US-09-982-402-1036
; Sequence 1036, Application US/09982402
; GENERAL INFORMATION:
; APPLICANT: [list inventors here]
; TITLE OF INVENTION: [list title here]
; FILE REFERENCE: [list docket ID here]
; CURRENT APPLICATION NUMBER: US/09/982,402
; CURRENT FILING DATE: 2001-10-17
; NUMBER OF SEQ ID NOS: 3553
; SOFTWARE: PERL Program
; SEQ ID NO 1036

```

```
LENGTH: 267
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No: 700196524H1
US-09-982-402-1036
```

```
Query Match          76.4%; Score 16.8; DB 6; Length 267;
Best Local Similarity 90.0%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 2 gtggatcacagaatttgat 21
||||| ||||||| |||
Db 180 gtggatgacagaattcgat 199
```

```
RESULT 7
US-09-849-526A-27340/c
Sequence 27340, Application US/09849526A
GENERAL INFORMATION:
APPLICANT: Ruff, Thomas G.
```

```
APPLICANT: Shukla, Hridayabhimanjan
TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
```

```
TITLE OF INVENTION: Plants
```

```
FILE REFERENCE: 16517,250138-21(51930)B]
```

```
CURRENT APPLICATION NUMBER: US/09/849,526A
```

```
PRIOR FILING DATE: 2001-05-07
```

```
PRIOR APPLICATION NUMBER: US 60/202,214
```

```
PRIOR FILING DATE: 2000-05-08
```

```
PRIOR APPLICATION NUMBER: US 09/654,617
```

```
PRIOR FILING DATE: 2000-09-05
```

```
PRIOR APPLICATION NUMBER: US 09/684,016
```

```
PRIOR FILING DATE: 2000-10-10
```

```
PRIOR APPLICATION NUMBER: US 09/816,660
```

```
PRIOR FILING DATE: 2001-03-26
```

```
NUMBER OF SEQ ID NOS: 30131
```

```
SEQ ID NO 27340
```

```
LENGTH: 551
```

```
TYPE: DNA
```

```
ORGANISM: Zea mays subsp. mexicana
```

```
FEATURE:
```

```
OTHER INFORMATION: Clone ID: uc-zmflteosinteg0d04b1
```

```
US-09-849-526A-27340
```

```
Query Match          76.4%; Score 16.8; DB 6; Length 551;
Best Local Similarity 90.0%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 3 ttgatcacagaatttgata 22
||||| ||||||| |||
Db 264 ttgcagcacaagatttgata 245
```

```
RESULT 8
US-09-815-264-26466/c
Sequence 26466, Application US/09815264
GENERAL INFORMATION:
```

```
APPLICANT: Boukharov, Andrey A.
```

```
APPLICANT: Cao, Yongwei
```

```
APPLICANT: Dotson, Stanton B.
```

```
APPLICANT: Koshi, Jeffrey M.
```

```
APPLICANT: Kovalic, David K.
```

```
APPLICANT: Liu, Jingdong
```

```
APPLICANT: McIninch, James
```

```
APPLICANT: Wu, Wei
```

```
TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof
```

```
FILE REFERENCE: 38-21(51237)G
```

```
CURRENT APPLICATION NUMBER: US/09/815,264
```

```
CURRENT FILING DATE: 2001-03-23
```

```
PRIOR APPLICATION NUMBER: US 09/620,392
```

```
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: US 09/702,134
PRIOR FILING DATE: 2000-10-31
```

```
NUMBER OF SEQ ID NOS: 109669
```

```
SEQ ID NO 26466
```

```
LENGTH: 849
```

```
TYPE: DNA
```

```
ORGANISM: Oryza sativa
```

```
US-09-815-264-26466
```

```
Query Match          76.4%; Score 16.8; DB 6; Length 849;
Best Local Similarity 90.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 2 gtggatcacagaatttgat 21
||||| ||||||| |||
Db 256 gtggatcactagatttgat 237
```

```
RESULT 9
US-09-815-264-63319
Sequence 63319, Application US/09815264
GENERAL INFORMATION:
```

```
APPLICANT: Boukharov, Andrey A.
```

```
APPLICANT: Cao, Yongwei
```

```
APPLICANT: Dotson, Stanton B.
```

```
APPLICANT: Koshi, Jeffrey M.
```

```
APPLICANT: Kovalic, David K.
```

```
APPLICANT: Liu, Jingdong
```

```
APPLICANT: McIninch, James
```

```
APPLICANT: Wu, Wei
```

```
TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof
```

```
FILE REFERENCE: 38-21(51237)G
```

```
CURRENT APPLICATION NUMBER: US/09/815,264
```

```
PRIOR FILING DATE: 2001-03-23
```

```
PRIOR APPLICATION NUMBER: US 09/620,392
```

```
PRIOR FILING DATE: 2000-07-19
```

```
PRIOR APPLICATION NUMBER: US 09/702,134
```

```
PRIOR FILING DATE: 2000-10-31
```

```
NUMBER OF SEQ ID NOS: 109669
```

```
SEQ ID NO 63319
```

```
LENGTH: 1046
```

```
TYPE: DNA
```

```
ORGANISM: Oryza sativa
```

```
US-09-815-264-63319
```

```
Query Match          76.4%; Score 16.8; DB 6; Length 1046;
Best Local Similarity 90.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 2 gtggatcacagaatttgat 21
||||| ||||||| |||
Db 791 gtggatcactagatttgat 810
```

```
RESULT 10
US-09-815-264-91672
Sequence 91672, Application US/09815264
GENERAL INFORMATION:
```

```
APPLICANT: Boukharov, Andrey A.
```

```
APPLICANT: Cao, Yongwei
```

```
APPLICANT: Dotson, Stanton B.
```

```
APPLICANT: Koshi, Jeffrey M.
```

```
APPLICANT: Kovalic, David K.
```

```
APPLICANT: Liu, Jingdong
```

```
APPLICANT: McIninch, James
```

```
APPLICANT: Wu, Wei
```

```
TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof
```

```
FILE REFERENCE: 38-21(51237)G
```

```
CURRENT APPLICATION NUMBER: US/09/815,264
```

```
CURRENT FILING DATE: 2001-03-23
```

```

; PRIOR APPLICATION NUMBER: US 09/620,392
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/702,134
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 109669
; SEQ ID NO 91672
; LENGTH: 1246
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-815-264-91672

Query Match          76.4%; Score 16.8; DB 6; Length 1246;
Best Local Similarity 90.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 gttgatacacaagattgat 21
||||| ||||| ||||| |||
Db 1013 gttgatacactagattgat 1032

RESULT 11
US-09-815-264-23140
; Sequence 23140, Application US/09815264
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Cao, Yongwei
; APPLICANT: Dotson, Stanton B.
; APPLICANT: Koshi, Jeffrey M.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: McIninch, James
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51237)G
; CURRENT APPLICATION NUMBER: US/09/815,264
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 09/620,392
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/702,134
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 109669
; SEQ ID NO 23140
; LENGTH: 1501
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-815-264-23140

Query Match          76.4%; Score 16.8; DB 6; Length 1501;
Best Local Similarity 90.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 ttgatacacaagattgata 22
||||| || ||||| |||||
Db 674 ttgatacacaagattgata 693

RESULT 12
US-09-815-264-46882
; Sequence 46882, Application US/09815264
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Cao, Yongwei
; APPLICANT: Dotson, Stanton B.
; APPLICANT: Koshi, Jeffrey M.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: McIninch, James
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51237)G
; CURRENT APPLICATION NUMBER: US/09/815,264

```

```

; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 09/620,392
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/702,134
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 109669
; SEQ ID NO 46882
; LENGTH: 1501
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1501)
; OTHER INFORMATION: unsure at all n locations
US-09-815-264-46882

Query Match          76.4%; Score 16.8; DB 6; Length 1501;
Best Local Similarity 90.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 ttgatacacaagattgata 22
||||| || ||||| |||||
Db 674 ttgatacacaagattgata 693

RESULT 13
US-09-815-264-53250
; Sequence 53250, Application US/09815264
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Cao, Yongwei
; APPLICANT: Dotson, Stanton B.
; APPLICANT: Koshi, Jeffrey M.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: McIninch, James
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51237)G
; CURRENT APPLICATION NUMBER: US/09/815,264
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 09/620,392
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/702,134
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 109669
; SEQ ID NO 53250
; LENGTH: 1501
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-815-264-53250

Query Match          76.4%; Score 16.8; DB 6; Length 1501;
Best Local Similarity 90.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 gttgatacacaagattgat 21
||||| ||||| ||||| |||
Db 88 gttgatacactagattgat 107

RESULT 14
US-09-815-264-76889
; Sequence 76889, Application US/09815264
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Cao, Yongwei
; APPLICANT: Dotson, Stanton B.
; APPLICANT: Koshi, Jeffrey M.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong

```



```
; APPLICANT: McIninch, James
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51237)G
; CURRENT APPLICATION NUMBER: US/09/815,264
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 09/620,392
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/702,134
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 109669
; SEQ ID NO 76889
; LENGTH: 6940
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(6940)
; OTHER INFORMATION: unsure at all n locations
US-09-815-264-76889
```

```
Query Match 76.4%; Score 16.8; DB 6; Length 6940;
Best Local Similarity 90.0%; Pred. No. 1.7e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 3 ttggatcacagaatttgata 22
||||| || ||||| |||||
Db 3165 ttggaccacaaagatttgata 3184
```

```
RESULT 15
US-09-815-264-81045/C
; Sequence 81045, Application US/09815264
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Cao, Yongwei
; APPLICANT: Dotson, Stanton B.
; APPLICANT: Koshl, Jeffrey M.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingtong
; APPLICANT: McIninch, James
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: Genomic Plant Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51237)G
; CURRENT APPLICATION NUMBER: US/09/815,264
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 09/620,392
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US 09/702,134
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 109669
; SEQ ID NO 81045
; LENGTH: 12081
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-815-264-81045
```

```
Query Match 76.4%; Score 16.8; DB 6; Length 12081;
Best Local Similarity 90.0%; Pred. No. 1.7e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 2 gttgatacacaagatttgat 21
||||| ||||| ||||| |||
Db 4391 GTTGATCACTAGATTAGAT 4372
```

Search completed: December 15, 2001, 05:58:19  
Job time: 12262 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 05:42:55 ; Search time 7962.77 seconds  
(without alignments)  
15.407 Million cell updates/sec

Title: US-09-380-826a-4  
Perfect score: 7  
Sequence: 1 tgttga 7

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 17159718 seqs, 876320856 residues

Total number of hits satisfying chosen parameters: 34319436

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 10%

Listing first 45 summaries

Database : Pending\_Patents\_Nr\_Main:\*

```
1: /cgn2_6/pdata/2/pna/US0611.COMB.seq:*
2: /cgn2_6/pdata/2/pna/US0612.COMB.seq:*
3: /cgn2_6/pdata/2/pna/US0613.COMB.seq:*
4: /cgn2_6/pdata/2/pna/US0614.COMB.seq:*
5: /cgn2_6/pdata/2/pna/US0615.COMB.seq:*
6: /cgn2_6/pdata/2/pna/US0616.COMB.seq:*
7: /cgn2_6/pdata/2/pna/US0617.COMB.seq:*
8: /cgn2_6/pdata/2/pna/US0618.COMB.seq:*
9: /cgn2_6/pdata/2/pna/US0619.COMB.seq:*
10: /cgn2_6/pdata/2/pna/US0620.COMB.seq:*
11: /cgn2_6/pdata/2/pna/US0621.COMB.seq:*
12: /cgn2_6/pdata/2/pna/US0622.COMB.seq:*
13: /cgn2_6/pdata/2/pna/US0623.COMB.seq:*
14: /cgn2_6/pdata/2/pna/US0624.COMB.seq:*
15: /cgn2_6/pdata/2/pna/US0625.COMB.seq:*
16: /cgn2_6/pdata/2/pna/US0626.COMB.seq:*
17: /cgn2_6/pdata/2/pna/US0627.COMB.seq:*
18: /cgn2_6/pdata/2/pna/US0628.COMB.seq:*
19: /cgn2_6/pdata/2/pna/US0629.COMB.seq:*
20: /cgn2_6/pdata/2/pna/US0630.COMB.seq:*
21: /cgn2_6/pdata/2/pna/US0631.COMB.seq:*
22: /cgn2_6/pdata/2/pna/US0632.COMB.seq:*
23: /cgn2_6/pdata/2/pna/US0633.COMB.seq:*
24: /cgn2_6/pdata/2/pna/US0634.COMB.seq:*
25: /cgn2_6/pdata/2/pna/US0635.COMB.seq:*
26: /cgn2_6/pdata/2/pna/US0636.COMB.seq:*
27: /cgn2_6/pdata/2/pna/US0637.COMB.seq:*
28: /cgn2_6/pdata/2/pna/US0638.COMB.seq:*
29: /cgn2_6/pdata/2/pna/US0639.COMB.seq:*
30: /cgn2_6/pdata/2/pna/US0640.COMB.seq:*
31: /cgn2_6/pdata/2/pna/US0641.COMB.seq:*
32: /cgn2_6/pdata/2/pna/US0642.COMB.seq:*
33: /cgn2_6/pdata/2/pna/US0643.COMB.seq:*
34: /cgn2_6/pdata/2/pna/US0644.COMB.seq:*
35: /cgn2_6/pdata/2/pna/US0645.COMB.seq:*
36: /cgn2_6/pdata/2/pna/US0646.COMB.seq:*
37: /cgn2_6/pdata/2/pna/US0647.COMB.seq:*
38: /cgn2_6/pdata/2/pna/US0648.COMB.seq:*
39: /cgn2_6/pdata/2/pna/US0649.COMB.seq:*
40: /cgn2_6/pdata/2/pna/US0650.COMB.seq:*
41: /cgn2_6/pdata/2/pna/US0651.COMB.seq:*
42: /cgn2_6/pdata/2/pna/US0652.COMB.seq:*
43: /cgn2_6/pdata/2/pna/US0653.COMB.seq:*
```

```
44: /cgn2_6/pdata/2/pna/US0654.COMB.seq:*
45: /cgn2_6/pdata/2/pna/US0655.COMB.seq:*
46: /cgn2_6/pdata/2/pna/US0656.COMB.seq:*
47: /cgn2_6/pdata/2/pna/US0657.COMB.seq:*
48: /cgn2_6/pdata/2/pna/US0658.COMB.seq:*
49: /cgn2_6/pdata/2/pna/US0659.COMB.seq:*
50: /cgn2_6/pdata/2/pna/US0660.COMB.seq:*
51: /cgn2_6/pdata/2/pna/US0661.COMB.seq:*
52: /cgn2_6/pdata/2/pna/US0662.COMB.seq:*
53: /cgn2_6/pdata/2/pna/US0663.COMB.seq:*
54: /cgn2_6/pdata/2/pna/US0664.COMB.seq:*
55: /cgn2_6/pdata/2/pna/US0665.COMB.seq:*
56: /cgn2_6/pdata/2/pna/US0666.COMB.seq:*
57: /cgn2_6/pdata/2/pna/US0667.COMB.seq:*
58: /cgn2_6/pdata/2/pna/US0668.COMB.seq:*
59: /cgn2_6/pdata/2/pna/US0669.COMB.seq:*
60: /cgn2_6/pdata/2/pna/US0670.COMB.seq:*
61: /cgn2_6/pdata/2/pna/US0671.COMB.seq:*
62: /cgn2_6/pdata/2/pna/US0672.COMB.seq:*
63: /cgn2_6/pdata/2/pna/US0673.COMB.seq:*
64: /cgn2_6/pdata/2/pna/US0674.COMB.seq:*
65: /cgn2_6/pdata/2/pna/US0675.COMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7	100.0	7	17 US-09-380-826a-4	Sequence 44, Appl
2	7	100.0	9	14 US-09-065-775-44	Sequence 44, Appl
3	7	100.0	32	US-09-916-466-170	Sequence 170, App
4	7	100.0	10	1 PCT-US00-24784-149	Sequence 149, App
5	7	100.0	10	1 PCT-US01-03620-47	Sequence 47, Appl
6	7	100.0	10	1 PCT-US01-18811A-45	Sequence 45, Appl
7	7	100.0	10	1 PCT-US94-12882-13	Sequence 13, Appl
8	7	100.0	10	1 PCT-US99-13800-679	Sequence 679, App
9	7	100.0	5	US-08-141-969-13	Sequence 13, Appl
10	7	100.0	10	16 US-09-213-932A-113	Sequence 13, Appl
11	7	100.0	10	17 US-09-335-032-2036	Sequence 2036, Ap
12	7	100.0	10	17 US-09-335-032-8949	Sequence 8949, Ap
13	7	100.0	10	17 US-09-335-032-11545	Sequence 11545, A
14	7	100.0	10	17 US-09-336-376-1209	Sequence 1209, Ap
15	7	100.0	10	17 US-09-336-376-4002	Sequence 4002, Ap
16	7	100.0	10	17 US-09-336-376-4328	Sequence 4328, Ap
17	7	100.0	10	17 US-09-336-376-5236	Sequence 5236, Ap
18	7	100.0	10	17 US-09-535-088-629	Sequence 629, Appl
19	7	100.0	10	23 US-09-601-844-18	Sequence 18, Appl
20	7	100.0	10	32 US-09-916-466-128	Sequence 128, Appl
21	7	100.0	11	32 US-09-945-505-30	Sequence 30, Appl
22	7	100.0	11	32 US-09-916-466-96	Sequence 96, Appl
23	7	100.0	12	1 PCT-US98-26935-386	Sequence 386, App
24	7	100.0	12	16 US-09-215-436-386	Sequence 386, App
25	7	100.0	13	19 US-09-528-209A-2234	Sequence 2234, Ap
26	7	100.0	13	19 US-09-528-209A-3083	Sequence 3083, Ap
27	7	100.0	13	19 US-09-528-209A-7388	Sequence 7388, Ap
28	7	100.0	12	32 US-09-586-479-2	Sequence 2, Appl
29	7	100.0	12	32 US-09-916-466-87	Sequence 87, Appl
30	7	100.0	13	1 PCT-US98-26935-274	Sequence 274, App
31	7	100.0	13	16 US-09-215-436-274	Sequence 274, App
32	7	100.0	13	19 US-09-528-209A-1123	Sequence 1123, Ap
33	7	100.0	14	15 US-09-164-961-1123	Sequence 1123, Ap
34	7	100.0	14	16 US-09-274-553-2956	Sequence 2956, Ap
35	7	100.0	14	16 US-09-274-553B-1407	Sequence 1407, Ap
36	7	100.0	14	16 US-09-274-553C-1407	Sequence 1407, Ap
37	7	100.0	14	16 US-09-277-026B-4887	Sequence 4887, Ap

c 42	7	100.0	14	17	US-09-341-700A-419	Sequence 419, App
c 43	7	100.0	14	17	US-09-341-700A-421	Sequence 421, App
c 44	7	100.0	14	17	US-09-341-700A-1193	Sequence 1193, App
c 45	7	100.0	14	19	US-09-504-231A-1407	Sequence 1407, App

## ALIGNMENTS

RESULT 1  
US-09-380-826A-4  
; Sequence 4, Application US/09380826A  
; GENERAL INFORMATION:  
; APPLICANT: Chappel, Rod  
; TITLE OF INVENTION: LEPTOSPIRA PATHOGENS  
; FILE REFERENCE: DAVIE79.001APC  
; CURRENT APPLICATION NUMBER: US/09/380, 826A  
; CURRENT FILING DATE: 1999-11-22  
; PRIOR APPLICATION NUMBER: PCT/AU98/00145  
; PRIOR FILING DATE: 1998-03-06  
; PRIOR APPLICATION NUMBER: AU P05494/97  
; PRIOR FILING DATE: 1997-03-07  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 7  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Primer for L. faeni.  
US-09-380-826A-4

Query Match  
Best Local Similarity 100.0%; Score 7; DB 17; Length 7;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7  
|111111  
DB 1 tgttga 7

RESULT 2  
US-09-065-775-44  
; Sequence 44, Application US/09065775A  
; GENERAL INFORMATION:  
; APPLICANT: Cen, Hui  
; TITLE OF INVENTION: Function-Based Gene Discovery  
; FILE REFERENCE: 59818  
; CURRENT APPLICATION NUMBER: US/09/065, 775A  
; CURRENT FILING DATE: 1998-04-24  
; NUMBER OF SEQ ID NOS: 105  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 44  
; LENGTH: 9  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-09-065-775-44

Query Match  
Best Local Similarity 100.0%; Score 7; DB 14; Length 9;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7  
|111111  
DB 3 tgttga 9

RESULT 3  
US-09-916-466-170  
; Sequence 170, Application US/09916466  
; GENERAL INFORMATION:  
; APPLICANT: Ribozyne Pharmaceuticals, Inc.  
; APPLICANT: Akhtar, Saghir  
; APPLICANT: McSwiggen, Jim  
; TITLE OF INVENTION: Enzymatic Nucleic Acid Treatment of Diseases or conditions Rel  
; FILE REFERENCE: MBH00-958-J (400/032)  
; CURRENT APPLICATION NUMBER: US/09/916,466  
; CURRENT FILING DATE: 2001-07-25  
; NUMBER OF SEQ ID NOS: 446  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 170  
; LENGTH: 9  
; TYPE: RNA  
; ORGANISM: Homo sapiens  
US-09-916-466-170

Query Match  
Best Local Similarity 57.1%; Score 7; DB 32; Length 9;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7  
|111111  
DB 1 ugnugga 7

RESULT 4  
PCT-US00-24784-149  
; Sequence 149, Application PC/TUS0024784  
; GENERAL INFORMATION:  
; APPLICANT: Nandabalan, Krishnan  
; APPLICANT: Stephens, Joel Claiborne  
; APPLICANT: Denton, R. Rex  
; APPLICANT: Chew, Anne  
; APPLICANT: Choi, Julie Y.  
; TITLE OF INVENTION: DRUG TARGET ISOGENES:  
; TITLE OF INVENTION: POLYMORPHISMS IN THE UNCOUPLING PROTEIN 3  
; FILE REFERENCE: UCP3-0009PCT  
; CURRENT APPLICATION NUMBER: PCT/US00/24784  
; CURRENT FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: MMH009  
; PRIOR FILING DATE: 1999-09-08  
; NUMBER OF SEQ ID NOS: 190  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 149  
; LENGTH: 10  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
PCT-US00-24784-149

Query Match  
Best Local Similarity 100.0%; Score 7; DB 1; Length 10;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7  
|111111  
DB 1 tgttga 7

RESULT 5  
PCT-US01-03620-47  
; Sequence 47, Application PC/TUS0103620  
; GENERAL INFORMATION:  
; APPLICANT: Genaisance Pharmaceuticals, Inc.  
; APPLICANT: Chew, Anne  
; APPLICANT: Koshiy, Beena  
; APPLICANT: Choi, Julie Y.

APPLICANT: Stephens, J. Claiborne  
TITLE OF INVENTION: DRUG TARGET ISOGENES: POLYMORPHISMS IN THE ANGIOTENSIN  
RECEPTOR 2 GENE  
FILE REFERENCE: MMH-0045 AGTR2  
CURRENT APPLICATION NUMBER: PCT/US01/03620  
CURRENT FILING DATE: 2001-02-02  
PRIOR APPLICATION NUMBER: 60/179,922  
PRIOR FILING DATE: 2000-02-03  
NUMBER OF SEQ ID NOS: 74  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 47  
LENGTH: 10  
TYPE: DNA  
ORGANISM: Homo sapiens  
PCT-US01-03620-47

Query Match 100.0%; Score 7; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 4.9e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 tgttgcga 7  
|111111|  
Db 2 tgttgcga 8

RESULT 6  
PCT-US01-18811A-45  
Sequence 45; Application PC/TUS0118811A  
GENERAL INFORMATION:  
APPLICANT: Genesance Pharmaceuticals, Inc.  
APPLICANT: Bentivegna, Steven C.  
APPLICANT: Bieglecki, Karyn  
APPLICANT: Duda, Amy  
APPLICANT: Kazeml, Amir  
APPLICANT: Koshy, Beena  
TITLE OF INVENTION: Haplotypes of the BDKRB2 Gene  
FILE REFERENCE: MMH-0740PCT BDKRB2  
CURRENT APPLICATION NUMBER: PCT/US01/18811A  
CURRENT FILING DATE: 2001-06-11  
PRIOR APPLICATION NUMBER: 60/210,575  
PRIOR FILING DATE: 2000-06-09  
NUMBER OF SEQ ID NOS: 57  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 45  
LENGTH: 10  
TYPE: DNA  
ORGANISM: Homo sapiens  
PCT-US01-18811A-45

Query Match 100.0%; Score 7; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 4.9e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 tgttcga 7  
|111111|  
Db 3 tgttcga 9

RESULT 7  
PCT-US94-12282-13/C  
Sequence 13; Application PC/TUS9412282  
GENERAL INFORMATION:  
APPLICANT: Beattie, Kenneth L.  
TITLE OF INVENTION: Microfabricated, Flowthrough Porous  
Apparatus for Discrete Detection of Binding Reactions  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Vinson & Elkins  
STREET: 1455 Pennsylvania Avenue, N.W.  
CITY: Washington  
STATE: D.C.

COUNTRY: U.S.A.  
ZIP: 20004-1008  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/12282  
FILING DATE: 27-Oct-1994  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/141,969  
FILING DATE: 28-Oct-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Sanzo, Michael A.  
REGISTRATION NUMBER: 36,912  
REFERENCE/DOCKET NUMBER: HARCO001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)639-6500  
TELEFAX: (202)639-6604  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: both  
TOPOLOGY: linear  
PCT-US94-12282-13

Query Match 100.0%; Score 7; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 4.9e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 tgttcga 7  
|111111|  
Db 7 tgttcga 1

RESULT 8  
PCT-US99-13800-679  
Sequence 679; Application PC/TUS9913800  
GENERAL INFORMATION:  
APPLICANT: Genzyme Corporation  
APPLICANT: Roberts, Bruce L.  
APPLICANT: Shankar, Srinivas  
TITLE OF INVENTION: PREPARATION AND USE OF SUPERIOR VACCINES  
FILE REFERENCE: 68126881206940  
CURRENT APPLICATION NUMBER: PCT/US99/13800  
CURRENT FILING DATE: 1999-06-18  
EARLIER APPLICATION NUMBER: 60/090,039  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/090,040  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/090,041  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/089,853  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/089,997  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/090,079  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/090,035  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/089,993  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/089,992  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/090,072  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/089,878  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/089,991

EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/090,000  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/090,048  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/089,999  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/090,043  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/090,042  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/090,036  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/090,044  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/089,844  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/090,080  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/089,833  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/089,994  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/090,078  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/090,077  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/090,076  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/090,045  
EARLIER FILING DATE: 1998-06-19  
EARLIER APPLICATION NUMBER: 60/111,715  
EARLIER FILING DATE: 1998-12-08  
NUMBER OF SEQ ID NOS: 2138  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO: 679  
LENGTH: 10  
TYPE: DNA  
ORGANISM: Homo sapiens  
PCT-US99-13800-679

Query Match 100.0%; Score 7; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 4.9e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tgttga 7  
1111111  
DB 3 tgttga 9

RESULT 9  
US-08-141-969-13/C  
Sequence 13, Application US/08141969  
GENERAL INFORMATION:  
APPLICANT: Beattie, Kenneth L.  
TITLE OF INVENTION: Microfabricated, Flowthrough Porous  
TITLE OF INVENTION: Apparatus for Discrete Detection of Binding Reactions  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Vinson & Elkins  
STREET: 1455 Pennsylvania Avenue, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20004-1008  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/141,969  
FILING DATE: 28-OCT-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Sanzo, Michael A.  
REGISTRATION NUMBER: 36,912  
REFERENCE/DOCKET NUMBER: HARC0001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)639-6500  
TELEFAX: (202)639-6604  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: both  
TOPOLOGY: linear  
US-08-141-969-13

Query Match 100.0%; Score 7; DB 5; Length 10;  
Best Local Similarity 100.0%; Pred. No. 4.9e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tgttga 7  
1111111  
DB 7 TGTGGA 1

RESULT 10  
US-09-213-932A-13/C  
Sequence 13, Application US/09213932A  
GENERAL INFORMATION:  
APPLICANT: Beattie, Kenneth L.  
TITLE OF INVENTION: Microfabricated, Flowthrough Porous  
TITLE OF INVENTION: Apparatus for Discrete Detection of Binding Reactions  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Vinson & Elkins  
STREET: 1455 Pennsylvania Avenue, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20004-1008  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/213,932A  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/631,751  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Sanzo, Michael A.  
REGISTRATION NUMBER: 36,912  
REFERENCE/DOCKET NUMBER: HARC0001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)639-6500  
TELEFAX: (202)639-6604  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: both  
TOPOLOGY: linear  
US-09-213-932A-13

Query Match 100.0%; Score 7; DB 16; Length 10;  
Best Local Similarity 100.0%; Pred. No. 4.9e+05;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7  
|111111  
Db 7 TGTGGA 1

RESULT 11  
; Sequence 2036, Application US/09335032  
; GENERAL INFORMATION:  
; APPLICANT: Velculescu, Victor  
; APPLICANT: Vogelstein, Bert  
; APPLICANT: Kinzler, Kenneth  
; TITLE OF INVENTION: Characterization of the Yeast  
; TITLE OF INVENTION: Transcriptome  
; FILE REFERENCE: 01107.78572  
; CURRENT APPLICATION NUMBER: US/09/335,032  
; CURRENT FILING DATE: 1999-06-16  
; PRIOR APPLICATION NUMBER: US 60/035,917  
; PRIOR FILING DATE: 1997-01-23  
; PRIOR APPLICATION NUMBER: US 09/012,031  
; PRIOR FILING DATE: 1998-01-22  
; NUMBER OF SEQ ID NOS: 12219  
; SOFTWARE: FastSeq for Windows, Version 4.0  
; SEQ ID NO 2036  
; LENGTH: 10  
; TYPE: DNA  
; ORGANISM: Saccharomyces cerevisiae  
US-09-335-032-2036

Query Match 100.0%; Score 7; DB 17; Length 10;  
Best Local Similarity 100.0%; Pred. No. 4.9e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7  
|111111  
Db 10 TGTGGA 4

RESULT 12  
US-09-335-032-8949/c  
; Sequence 8949, Application US/09335032  
; GENERAL INFORMATION:  
; APPLICANT: Velculescu, Victor  
; APPLICANT: Vogelstein, Bert  
; APPLICANT: Kinzler, Kenneth  
; TITLE OF INVENTION: Characterization of the Yeast  
; TITLE OF INVENTION: Transcriptome  
; FILE REFERENCE: 01107.78572  
; CURRENT APPLICATION NUMBER: US/09/335,032  
; CURRENT FILING DATE: 1999-06-16  
; PRIOR APPLICATION NUMBER: US 60/035,917  
; PRIOR FILING DATE: 1997-01-23  
; PRIOR APPLICATION NUMBER: US 09/012,031  
; PRIOR FILING DATE: 1998-01-22  
; NUMBER OF SEQ ID NOS: 12219  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8949  
; LENGTH: 10  
; TYPE: DNA  
; ORGANISM: Saccharomyces cerevisiae  
US-09-335-032-8949

Query Match 100.0%; Score 7; DB 17; Length 10;  
Best Local Similarity 100.0%; Pred. No. 4.9e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7  
|111111  
Db 10 TGTGGA 4

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 13  
US-09-335-032-11545  
; Sequence 11545, Application US/09335032  
; GENERAL INFORMATION:  
; APPLICANT: Velculescu, Victor  
; APPLICANT: Vogelstein, Bert  
; APPLICANT: Kinzler, Kenneth  
; TITLE OF INVENTION: Characterization of the Yeast  
; TITLE OF INVENTION: Transcriptome  
; FILE REFERENCE: 01107.78572  
; CURRENT APPLICATION NUMBER: US/09/335,032  
; CURRENT FILING DATE: 1999-06-16  
; PRIOR APPLICATION NUMBER: US 60/035,917  
; PRIOR FILING DATE: 1997-01-23  
; PRIOR APPLICATION NUMBER: US 09/012,031  
; PRIOR FILING DATE: 1998-01-22  
; NUMBER OF SEQ ID NOS: 12219  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11545  
; LENGTH: 10  
; TYPE: DNA  
; ORGANISM: Saccharomyces cerevisiae  
US-09-335-032-11545

Query Match 100.0%; Score 7; DB 17; Length 10;  
Best Local Similarity 100.0%; Pred. No. 4.9e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7  
|111111  
Db 1 tgttga 7

RESULT 14  
US-09-336-376-1209  
; Sequence 1209, Application US/09336376  
; GENERAL INFORMATION:  
; APPLICANT: Roberts, Bruce L.  
; TITLE OF INVENTION: POLYNUCLEOTIDE POPULATION ISOLATED FROM  
; TITLE OF INVENTION: HUMAN MONOCYTES  
; FILE REFERENCE: 126745205600  
; CURRENT APPLICATION NUMBER: US/09/336,376  
; CURRENT FILING DATE: 1999-06-17  
; EARLIER APPLICATION NUMBER: 60/090,039  
; EARLIER FILING DATE: 1998-06-19  
; EARLIER APPLICATION NUMBER: 60/090,040  
; EARLIER FILING DATE: 1998-06-19  
; EARLIER APPLICATION NUMBER: 60/090,041  
; EARLIER FILING DATE: 1998-06-19  
; EARLIER APPLICATION NUMBER: 60/089,853  
; EARLIER FILING DATE: 1998-06-19  
; EARLIER APPLICATION NUMBER: 60/089,997  
; EARLIER FILING DATE: 1998-06-19  
; EARLIER APPLICATION NUMBER: 60/090,079  
; EARLIER FILING DATE: 1998-06-19  
; EARLIER APPLICATION NUMBER: 60/090,035  
; EARLIER FILING DATE: 1998-06-19  
; EARLIER APPLICATION NUMBER: 60/089,993  
; EARLIER FILING DATE: 1998-06-19  
; EARLIER APPLICATION NUMBER: 60/089,992  
; EARLIER FILING DATE: 1998-06-19  
; EARLIER APPLICATION NUMBER: 60/090,072  
; EARLIER FILING DATE: 1998-06-19  
; EARLIER APPLICATION NUMBER: 60/089,878  
; EARLIER FILING DATE: 1998-06-19  
; EARLIER APPLICATION NUMBER: 60/089,991  
; EARLIER FILING DATE: 1998-06-19  
; EARLIER APPLICATION NUMBER: 60/090,000  
; EARLIER FILING DATE: 1998-06-19  
; EARLIER APPLICATION NUMBER: 60/090,048

```
;; EARLIER FILING DATE: 1998-06-19
;; EARLIER APPLICATION NUMBER: 60/089,999
;; EARLIER FILING DATE: 1998-06-19
;; EARLIER APPLICATION NUMBER: 60/090,043
;; EARLIER FILING DATE: 1998-06-19
;; EARLIER APPLICATION NUMBER: 60/090,042
;; EARLIER FILING DATE: 1998-06-19
;; EARLIER APPLICATION NUMBER: 60/090,036
;; EARLIER FILING DATE: 1998-06-19
;; EARLIER APPLICATION NUMBER: 60/090,044
;; EARLIER FILING DATE: 1998-06-19
;; EARLIER APPLICATION NUMBER: 60/089,844
;; EARLIER FILING DATE: 1998-06-19
;; EARLIER APPLICATION NUMBER: 60/089,833
;; EARLIER FILING DATE: 1998-06-19
;; EARLIER APPLICATION NUMBER: 60/089,994
;; EARLIER FILING DATE: 1998-06-19
;; EARLIER APPLICATION NUMBER: 60/090,077
;; EARLIER FILING DATE: 1998-06-19
;; EARLIER APPLICATION NUMBER: 60/090,078
;; EARLIER FILING DATE: 1998-06-19
;; EARLIER APPLICATION NUMBER: 60/090,047
;; EARLIER FILING DATE: 1998-06-19
;; EARLIER APPLICATION NUMBER: 60/090,076
;; EARLIER FILING DATE: 1998-06-19
;; EARLIER APPLICATION NUMBER: 60/090,045
;; EARLIER FILING DATE: 1998-06-19
;; EARLIER APPLICATION NUMBER: 60/111,715
;; EARLIER FILING DATE: 1998-12-08
;; NUMBER OF SEQ ID NOS: 5980
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO: 1209
;; LENGTH: 10
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-336-376-1209
```

```
Query Match          100.0%; Score 7; DB 17; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.9e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 tgttga 7
        |||||||
Db       4 tgttga 10
```

```
RESULT 15
US-09-336-376-4002
; Sequence 4002, Application us/09336376
; GENERAL INFORMATION:
; APPLICANT: Roberts, Bruce L.
; TITLE OF INVENTION: POLYNUCLEOTIDE POPULATION ISOLATED FROM
; FILE REFERENCE: 126745205600
; CURRENT APPLICATION NUMBER: us/09/336,376
; EARLIER FILING DATE: 1999-06-17
; EARLIER APPLICATION NUMBER: 60/090,039
; EARLIER FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/090,040
; EARLIER FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/090,041
; EARLIER FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/089,853
; EARLIER FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/089,997
; EARLIER FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/090,079
; EARLIER FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/090,035
; EARLIER FILING DATE: 1998-06-19
```

```
;; EARLIER APPLICATION NUMBER: 60/089,993
;; EARLIER FILING DATE: 1998-06-19
;; EARLIER APPLICATION NUMBER: 60/089,992
;; EARLIER FILING DATE: 1998-06-19
;; EARLIER APPLICATION NUMBER: 60/090,072
;; EARLIER FILING DATE: 1998-06-19
;; EARLIER APPLICATION NUMBER: 60/089,878
;; EARLIER FILING DATE: 1998-06-19
;; EARLIER APPLICATION NUMBER: 60/089,991
;; EARLIER FILING DATE: 1998-06-19
;; EARLIER APPLICATION NUMBER: 60/090,000
;; EARLIER FILING DATE: 1998-06-19
;; EARLIER APPLICATION NUMBER: 60/090,048
;; EARLIER FILING DATE: 1998-06-19
;; EARLIER APPLICATION NUMBER: 60/089,999
;; EARLIER FILING DATE: 1998-06-19
;; EARLIER APPLICATION NUMBER: 60/090,043
;; EARLIER FILING DATE: 1998-06-19
;; EARLIER APPLICATION NUMBER: 60/090,042
;; EARLIER FILING DATE: 1998-06-19
;; EARLIER APPLICATION NUMBER: 60/090,036
;; EARLIER FILING DATE: 1998-06-19
;; EARLIER APPLICATION NUMBER: 60/090,044
;; EARLIER FILING DATE: 1998-06-19
;; EARLIER APPLICATION NUMBER: 60/089,844
;; EARLIER FILING DATE: 1998-06-19
;; EARLIER APPLICATION NUMBER: 60/090,080
;; EARLIER FILING DATE: 1998-06-19
;; EARLIER APPLICATION NUMBER: 60/089,833
;; EARLIER FILING DATE: 1998-06-19
;; EARLIER APPLICATION NUMBER: 60/089,994
;; EARLIER FILING DATE: 1998-06-19
;; EARLIER APPLICATION NUMBER: 60/090,077
;; EARLIER FILING DATE: 1998-06-19
;; EARLIER APPLICATION NUMBER: 60/090,078
;; EARLIER FILING DATE: 1998-06-19
;; EARLIER APPLICATION NUMBER: 60/090,047
;; EARLIER FILING DATE: 1998-06-19
;; EARLIER APPLICATION NUMBER: 60/090,076
;; EARLIER FILING DATE: 1998-06-19
;; EARLIER APPLICATION NUMBER: 60/090,045
;; EARLIER FILING DATE: 1998-06-19
;; EARLIER APPLICATION NUMBER: 60/111,715
;; EARLIER FILING DATE: 1998-12-08
;; NUMBER OF SEQ ID NOS: 5980
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO: 4002
;; LENGTH: 10
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-336-376-4002
```

```
Query Match          100.0%; Score 7; DB 17; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.9e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 tgttga 7
        |||||||
Db       4 tgttga 10
```

```
Search completed: December 15, 2001, 05:42:56
Job time: 15821 sec
```





**THIS PAGE BLANK (USPTO)**



```

; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 09/578,223
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 09/476,387
; PRIOR FILING DATE: 1999-12-30
; PRIOR APPLICATION NUMBER: 09/474,432
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 09/501,511
; PRIOR FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: 09/186,675
; PRIOR FILING DATE: 1998-11-04
; PRIOR APPLICATION NUMBER: 60/083,727
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/064,866
; PRIOR FILING DATE: 1997-11-05
; NUMBER OF SEQ ID NOS: 1558
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 181
; LENGTH: 13
; TYPE: RNA
; ORGANISM: Homo sapiens
; US-09-825-805-181
```

```
Query Match          100.0%; Score 7; DB 6; Length 13;
Best Local Similarity 57.1%; Pred. No. 6.1e+04;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 tgttga 7
    |||||
Db 6 ugungga 12
```

```
RESULT 3
PCT-US01-30411-7
; Sequence 7, Application PC/TUS0130411
; GENERAL INFORMATION:
; APPLICANT: Genaisance Pharmaceuticals, Inc.
; APPLICANT: Choi, Julie Y
; APPLICANT: Kilem, Stefanie E
; APPLICANT: Koshy, Beena
; APPLICANT: Parks, Katie E
; APPLICANT: Stephens, J. Claiborne
; TITLE OF INVENTION: HAPLOTYPES OF THE TCF2 GENE
; FILE REFERENCE: TCF2_MWH-1439PCT
; CURRENT APPLICATION NUMBER: PCT/US01/30411
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/235,710
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US01-30411-7
```

```
Query Match          100.0%; Score 7; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 6.1e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 tgttga 7
    |||||
Db 9 tgttga 15
```

```
RESULT 4
PCT-US01-30411-17/c
; Sequence 17, Application PC/TUS0130411
; GENERAL INFORMATION:
; APPLICANT: Genaisance Pharmaceuticals, Inc.
; APPLICANT: Choi, Julie Y
```

```

; APPLICANT: Kilem, Stefanie E
; APPLICANT: Koshy, Beena
; APPLICANT: Parks, Katie E
; APPLICANT: Stephens, J. Claiborne
; TITLE OF INVENTION: HAPLOTYPES OF THE TCF2 GENE
; FILE REFERENCE: TCF2_MWH-1439PCT
; CURRENT APPLICATION NUMBER: PCT/US01/30411
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/235,710
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US01-30411-17
```

```
Query Match          100.0%; Score 7; DB 1; Length 15;
Best Local Similarity 57.1%; Pred. No. 6.1e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 tgttga 7
    |||||
Db 13 TGTGTGA 7
```

```
RESULT 5
US-09-825-805-182
; Sequence 182, Application US/09825805
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: Belgelman, Leo
; APPLICANT: Beaudry, Amber
; APPLICANT: Karpeisky, Alex
; APPLICANT: Adamic, Jasenka Matulic
; APPLICANT: Sweedler, Dave
; APPLICANT: Zinnen, Shawn
; TITLE OF INVENTION: Nucleotide Triphosphate and their Incorporation Into Oligonucle
; FILE REFERENCE: MBH800-831-F (400/009)
; CURRENT APPLICATION NUMBER: US/09/825,805
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 09/578,223
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 09/476,387
; PRIOR FILING DATE: 1999-12-30
; PRIOR APPLICATION NUMBER: 09/474,432
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 09/301,511
; PRIOR FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: 09/186,675
; PRIOR FILING DATE: 1998-11-04
; PRIOR APPLICATION NUMBER: 60/083,727
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/064,866
; PRIOR FILING DATE: 1997-11-05
; NUMBER OF SEQ ID NOS: 1558
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 182
; LENGTH: 15
; TYPE: RNA
; ORGANISM: Homo sapiens
US-09-825-805-182
```

```
Query Match          100.0%; Score 7; DB 6; Length 15;
Best Local Similarity 57.1%; Pred. No. 6.1e+04;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 tgttga 7
    |||||
Db 7 ugungga 13
```

RESULT 6  
US-09-527-030C-105  
; Sequence 105, Application US/09527030C  
; GENERAL INFORMATION:  
; APPLICANT: VAN DOORN, Leen-Jan et al.  
; TITLE OF INVENTION: Detection and identification of Human Papillomavirus by PCR and t  
; FILE REFERENCE: 3501-0101P  
; CURRENT APPLICATION NUMBER: US/09/527,030C  
; CURRENT FILING DATE: 2000-03-16  
; NUMBER OF SEQ ID NOS: 497  
; SOFTWARE: PatentIn version 3.0.  
; SEQ ID NO 105  
; LENGTH: 16  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Type specific probe derived from the Human Papillomavirus (HPV)  
US-09-527-030C-105

Query Match 100.0%; Score 7; DB 6; Length 16;  
Best Local Similarity 100.0%; Pred. No. 6.1e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7  
:|||||  
Db 6 tgttga 12

RESULT 7  
US-09-527-030C-181  
; Sequence 181, Application US/09527030C  
; GENERAL INFORMATION:  
; APPLICANT: VAN DOORN, Leen-Jan et al.  
; TITLE OF INVENTION: Detection and identification of Human Papillomavirus by PCR and t  
; FILE REFERENCE: 3501-0101P  
; CURRENT APPLICATION NUMBER: US/09/527,030C  
; CURRENT FILING DATE: 2000-03-16  
; NUMBER OF SEQ ID NOS: 497  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 181  
; LENGTH: 17  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic probe derived from the Human Papillomavirus (HPV)  
US-09-527-030C-181

Query Match 100.0%; Score 7; DB 6; Length 17;  
Best Local Similarity 100.0%; Pred. No. 6.1e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7  
:|||||  
Db 7 tgttga 13

RESULT 8  
US-09-685-664B-1636  
; Sequence 1636, Application US/09685664B  
; GENERAL INFORMATION:  
; APPLICANT: Ribozyme Pharmaceuticals, Inc.  
; APPLICANT: Payco, Pam  
; APPLICANT: McSwigen, Jim  
; APPLICANT: Stinchcomb, Dan  
; APPLICANT: Escobedo, Jaime  
; TITLE OF INVENTION: Method and Reagent for Treatment of Diseases or Conditions Relate  
; TITLE OF INVENTION: Levels of Vascular Endothelial Growth Factor Receptor

; FILE REFERENCE: MBH00-876-K (400/021)  
; CURRENT APPLICATION NUMBER: US/09/685,664B  
; CURRENT FILING DATE: 2000-10-10  
; PRIOR APPLICATION NUMBER: US 60/005,974  
; PRIOR FILING DATE: 1995-10-26  
; PRIOR APPLICATION NUMBER: US 08/584,040  
; PRIOR FILING DATE: 1996-01-08  
; PRIOR APPLICATION NUMBER: US 09/371,772  
; PRIOR FILING DATE: 1999-08-10  
; NUMBER OF SEQ ID NOS: 8231  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1636  
; LENGTH: 17  
; TYPE: RNA  
; ORGANISM: Homo sapiens  
US-09-685-664B-1636

Query Match 100.0%; Score 7; DB 6; Length 17;  
Best Local Similarity 57.1%; Pred. No. 6.1e+04;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7  
:|||||  
Db 10 uguuga 16

RESULT 9  
US-09-685-664B-1637  
; Sequence 1637, Application US/09685664B  
; GENERAL INFORMATION:  
; APPLICANT: Ribozyme Pharmaceuticals, Inc.  
; APPLICANT: Payco, Pam  
; APPLICANT: McSwigen, Jim  
; APPLICANT: Stinchcomb, Dan  
; APPLICANT: Escobedo, Jaime  
; TITLE OF INVENTION: Method and Reagent for Treatment of Diseases or Conditions Rel  
; FILE REFERENCE: MBH00-876-K (400/021)  
; CURRENT APPLICATION NUMBER: US/09/685,664B  
; CURRENT FILING DATE: 2000-10-10  
; PRIOR APPLICATION NUMBER: US 60/005,974  
; PRIOR FILING DATE: 1995-10-26  
; PRIOR APPLICATION NUMBER: US 08/584,040  
; PRIOR FILING DATE: 1996-01-08  
; PRIOR APPLICATION NUMBER: US 09/371,772  
; PRIOR FILING DATE: 1999-08-10  
; NUMBER OF SEQ ID NOS: 8231  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1637  
; LENGTH: 17  
; TYPE: RNA  
; ORGANISM: Homo sapiens  
US-09-685-664B-1637

Query Match 100.0%; Score 7; DB 6; Length 17;  
Best Local Similarity 57.1%; Pred. No. 6.1e+04;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7  
:|||||  
Db 6 uguuga 12

RESULT 10  
US-09-685-664B-3484  
; Sequence 3484, Application US/09685664B  
; GENERAL INFORMATION:  
; APPLICANT: Ribozyme Pharmaceuticals, Inc.  
; APPLICANT: Payco, Pam  
; APPLICANT: McSwigen, Jim  
; APPLICANT: Stinchcomb, Dan

APPLICANT: Escobedo, Jaime  
TITLE OF INVENTION: Method and Reagent for Treatment of Diseases or Conditions Related to Vascular Endothelial Growth Factor Receptor  
FILE REFERENCE: MBH00-876-K (400/021)  
CURRENT APPLICATION NUMBER: US/09/685,664B  
CURRENT FILING DATE: 2000-10-10  
PRIOR APPLICATION NUMBER: US 60/005,974  
PRIOR FILING DATE: 1995-10-26  
PRIOR APPLICATION NUMBER: US 08/584,040  
PRIOR FILING DATE: 1996-01-08  
PRIOR APPLICATION NUMBER: US 09/371,772  
PRIOR FILING DATE: 1999-08-10  
NUMBER OF SEQ ID NOS: 8231  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 3484  
LENGTH: 17  
TYPE: RNA  
ORGANISM: Mus musculus  
US-09-685-664B-3484

Query Match 100.0%; Score 7; DB 6; Length 17;  
Best Local Similarity 57.1%; Pred. No. 6.1e+04;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7  
:|||||  
DB 11 uguuga 17

RESULT 11  
US-09-685-664B-3485  
Sequence 3485, Application US/09685664B  
GENERAL INFORMATION:  
APPLICANT: Ribozyne Pharmaceuticals, Inc.  
APPLICANT: Pavco, Pam  
APPLICANT: McSwigen, Jim  
APPLICANT: Stinchcomb, Dan  
APPLICANT: Escobedo, Jaime  
TITLE OF INVENTION: Method and Reagent for Treatment of Diseases or Conditions Related to Vascular Endothelial Growth Factor Receptor  
FILE REFERENCE: MBH00-876-K (400/021)  
CURRENT APPLICATION NUMBER: US/09/685,664B  
CURRENT FILING DATE: 2000-10-10  
PRIOR APPLICATION NUMBER: US 60/005,974  
PRIOR FILING DATE: 1995-10-26  
PRIOR APPLICATION NUMBER: US 08/584,040  
PRIOR FILING DATE: 1996-01-08  
PRIOR APPLICATION NUMBER: US 09/371,772  
PRIOR FILING DATE: 1999-08-10  
NUMBER OF SEQ ID NOS: 8231  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 3485  
LENGTH: 17  
TYPE: RNA  
ORGANISM: Mus musculus  
US-09-685-664B-3485

Query Match 100.0%; Score 7; DB 6; Length 17;  
Best Local Similarity 57.1%; Pred. No. 6.1e+04;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7  
:|||||  
DB 6 uguuga 12

RESULT 12  
US-09-685-664B-3486  
Sequence 3486, Application US/09685664B  
GENERAL INFORMATION:  
APPLICANT: Ribozyne Pharmaceuticals, Inc.

APPLICANT: Pavco, Pam  
APPLICANT: McSwigen, Jim  
APPLICANT: Stinchcomb, Dan  
APPLICANT: Escobedo, Jaime  
TITLE OF INVENTION: Method and Reagent for Treatment of Diseases or Conditions Related to Vascular Endothelial Growth Factor Receptor  
FILE REFERENCE: MBH00-876-K (400/021)  
CURRENT APPLICATION NUMBER: US/09/685,664B  
CURRENT FILING DATE: 2000-10-10  
PRIOR APPLICATION NUMBER: US 60/005,974  
PRIOR FILING DATE: 1995-10-26  
PRIOR APPLICATION NUMBER: US 08/584,040  
PRIOR FILING DATE: 1996-01-08  
PRIOR APPLICATION NUMBER: US 09/371,772  
PRIOR FILING DATE: 1999-08-10  
NUMBER OF SEQ ID NOS: 8231  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 3486  
LENGTH: 17  
TYPE: RNA  
ORGANISM: Mus musculus  
US-09-685-664B-3486

Query Match 100.0%; Score 7; DB 6; Length 17;  
Best Local Similarity 57.1%; Pred. No. 6.1e+04;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7  
:|||||  
DB 11 uguuga 7

RESULT 13  
US-09-685-664B-3493/C  
Sequence 3493, Application US/09685664B  
GENERAL INFORMATION:  
APPLICANT: Ribozyne Pharmaceuticals, Inc.  
APPLICANT: Pavco, Pam  
APPLICANT: McSwigen, Jim  
APPLICANT: Stinchcomb, Dan  
APPLICANT: Escobedo, Jaime  
TITLE OF INVENTION: Method and Reagent for Treatment of Diseases or Conditions Related to Vascular Endothelial Growth Factor Receptor  
FILE REFERENCE: MBH00-876-K (400/021)  
CURRENT APPLICATION NUMBER: US/09/685,664B  
CURRENT FILING DATE: 2000-10-10  
PRIOR APPLICATION NUMBER: US 60/005,974  
PRIOR FILING DATE: 1995-10-26  
PRIOR APPLICATION NUMBER: US 08/584,040  
PRIOR FILING DATE: 1996-01-08  
PRIOR APPLICATION NUMBER: US 09/371,772  
PRIOR FILING DATE: 1999-08-10  
NUMBER OF SEQ ID NOS: 8231  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 3493  
LENGTH: 17  
TYPE: RNA  
ORGANISM: Mus musculus  
US-09-685-664B-3493

Query Match 100.0%; Score 7; DB 6; Length 17;  
Best Local Similarity 100.0%; Pred. No. 6.1e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgttga 7  
|||||||  
DB 14 tgttga 8

RESULT 14  
US-09-740-332-1100/C

```

; Sequence 1100, Application US/09740332
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals Inc.
; TITLE OF INVENTION: Enzymatic Nucleic Acid Treatment of Diseases or Conditions Relate
; TITLE OF INVENTION: Hepatitis C Virus Infection
; FILE REFERENCE: RPI 400/003
; CURRENT APPLICATION NUMBER: US/09/740,332
; CURRENT FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 9704
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1100
; LENGTH: 17
; TYPE: RNA
; ORGANISM: artificial sequence
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: oligonucleotide substrate
US-09-740-332-1100

```

```

Query Match          100.0%; Score 7; DB 6; Length 17;
Best Local Similarity 100.0%; Pred. No. 6.1e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 tgttga 7
        |||||
Db       16 TGTGGA 10

```

```

RESULT 15
US-09-740-332-1101/c
; Sequence 1101, Application US/09740332
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals Inc.
; TITLE OF INVENTION: Enzymatic Nucleic Acid Treatment of Diseases or Conditions Relate
; TITLE OF INVENTION: Hepatitis C Virus Infection
; FILE REFERENCE: RPI 400/003
; CURRENT APPLICATION NUMBER: US/09/740,332
; CURRENT FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 9704
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1101
; LENGTH: 17
; TYPE: RNA
; ORGANISM: artificial sequence
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: oligonucleotide substrate
US-09-740-332-1101

```

```

Query Match          100.0%; Score 7; DB 6; Length 17;
Best Local Similarity 100.0%; Pred. No. 6.1e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 tgttga 7
        |||||
Db       11 TGTGGA 5

```

Search completed: December 15, 2001, 05:58:20  
 Job time: 12263 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 01:11:50 ; Search time 2725.73 Seconds  
(without alignments)  
8939.379 Million cell updates/sec

Title: US-09-380-826A-1  
Perfect score: 1477  
Sequence: 1 gacatgcgcgcgaactaac.....ccgtaatgccttcctgcag 1477

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues  
Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:  
1: gb-ba:\*  
2: gb-hcg:\*  
3: gb-in:\*  
4: gb-om:\*  
5: gb-ov:\*  
6: gb-pat:\*  
7: gb-ph:\*  
8: gb-pl:\*  
9: gb-pr:\*  
10: gb-ro:\*  
11: gb-sts:\*  
12: gb-sy:\*  
13: gb-un:\*  
14: gb-vi:\*  
15: em-ba:\*  
16: em-fun:\*  
17: em-hum:\*  
18: em-in:\*  
19: em-om:\*  
20: em-or:\*  
21: em-ov:\*  
22: em-pat:\*  
23: em-ph:\*  
24: em-pl:\*  
25: em-ro:\*  
26: em-sts:\*  
27: em-sy:\*  
28: em-un:\*  
29: em-vi:\*  
30: em-htgo-hum:\*  
31: em-htgo-ity:\*  
32: em-htgo-rod:\*  
33: em-htgo-hum:\*  
34: em-htgo-ity:\*  
35: em-htgo-rod:\*  
36: em-htgo-other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

8

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1431.4	96.9	1481	1	LFU60594	U60594 Leptospira
2	1388.6	94.0	1450	1	LEA19243	Y19243 Leptospira
3	1295.6	87.7	1508	1	LIRRN16S	X17543 Leptospira
4	1270	86.0	1374	1	L116SRDNY	Z21634 L.inadai ge
5	1263.8	85.6	1483	1	MA4516S	U12676 Leptospira
6	1263.8	85.6	1489	1	MA5216S	U12677 Leptospira
7	1262.2	85.5	1513	1	LB012670	U12670 Leptospira
8	1261	85.4	1489	1	LB012669	U12669 Leptospira
9	1260.8	85.4	1516	1	LN012671	U12671 Leptospira
10	1254.6	84.9	1425	1	LB16SRDNP	U12630 L.borgpeter
11	1251.8	84.8	1486	1	WP4616S	U12637 Leptospira
12	1246.6	84.4	1419	1	LB16SRDNP	U12628 L.kirschner
13	1239.8	83.9	1494	1	LSU12672	U12672 Leptospira
14	1238.2	83.8	1418	1	LB16SRDNP	U12648 L.meyerl ge
15	1232.8	83.5	1396	1	L116SRRN	Z21817 L.interroga
16	1230	83.3	1415	1	LS16SRDNP	Z21649 L.santarosa
17	1225.4	83.0	1410	1	LB16SRDNP	Z21637 L.weilli ge
18	1216	82.3	1424	1	LN16SRDNP	Z21635 L.noquchil
19	1063.4	72.0	1523	1	LEPRR16SB	M88721 Leptospira
20	1031.8	65.9	1433	1	LB16SRDNP	Z98590 Leptospira
21	1027.6	66.6	1417	1	LB16SRDNP	Z21638 L.wolbachii
22	1022.2	69.2	1439	1	LB16SRDNP	Z98588 Leptospira
23	1003.6	67.9	1451	1	AA16SRDNP	Z21629 A.ancona ge
24	1003.4	67.9	1321	1	LEPRR16SA	M71241 Leptospira
25	1003.2	67.9	1366	1	AF167353	AF167353 Leptospir
26	999.8	67.7	1442	1	LB16SRDNP	Z98587 Leptospira
27	995.8	67.4	1410	1	CC16SRDNP	Z21631 C.canela ge
28	993	67.2	1398	1	LB16SRDNP	Z12821 L.biflex 1
29	991.8	67.1	1394	1	LB16SRDNP	Z26699 L.biflex 9
30	991.8	67.1	1394	1	LB16SRDNP	Z98592 Leptospira
31	991.6	67.1	1527	1	LEPRR16SC	M88719 Leptospira
32	991	67.1	1379	1	UJ16SRDNP	Z21633 J.-Jequilata
33	985.8	66.7	1427	1	LB16SRDNP	Z98593 Leptospira
34	977.8	66.2	1466	1	LB16SRDNP	Z98589 Leptospira
35	972	65.8	1445	1	LB16SRDNP	Z21632 L.illini ge
36	964.4	65.3	1413	1	LB16SRDNP	Z98591 Leptospira
37	915.8	62.0	1543	1	UFA19479	Z26970 L.unculture
38	884	52.9	1385	1	LB16SRDNP	Z26970 L.biflex 9
39	883.2	59.8	1475	1	UE081652	U81652 Unidentifie
40	874.4	59.2	1494	1	PA16SRRN	X77216 P.acidigall
41	866.6	58.7	1494	1	CG16SRRN	X75272 Clostridium
42	866.6	58.7	1525	1	MR17712	Y17712 Malonomas
43	864.2	58.5	1503	1	CT16SRNA	X72869 C.thermopol
44	857.8	58.1	1532	1	AB039334	AB039334 Brevibact
45	856.6	58.0	1379	1	LP16SRDNP	Z21636 L.parva gen

## ALIGNMENTS

RESULT	1	LOCUS	LFU60594	1481 bp	DNA	BCT	10-SEP-1998
DEFINITION		Leptospira	fainei 16S ribosomal RNA gene, partial sequence.				
ACCESSION		U60594					
VERSION		U60594.1	GI:1408219				
KEYWORDS							
SOURCE							
ORGANISM							
REFERENCE							
AUTHORS							
TITLE							
JOURNAL							
MEDLINE							
REFERENCE							
AUTHORS							

1 (bases 1 to 1481)  
Bacteria: Spirochaetales; Leptospiraceae; Leptospira.  
Leptospira fainei.  
Leptospira fainei.  
Perolat,P., Chappel,R.J., Adler,B., Baranton,G., Bulach,D.M., Billinghamst,M.L., Letocart,M., Merlen,F. and Serrano,M.S.  
Leptospira fainei sp. nov., isolated from pigs in Australia  
Int. J. Syst. Bacteriol. 48 Pt 3, 851-858 (1998)  
2 (bases 1 to 1481)  
Adler,B., Chappel,R.J., Baranton,G., Bulach,D.M., Billinghamst,M.L., Letocart,M., Merlen,F., Serrano,M.S. and Perolat,P.

TITLE		Direct Submission	
JOURNML		Submitted (12-JUN-1996) Microbiology, Monash University, Wellington	
RD., Claydon, VIC 3168, Australia			
FEATURES		Location/Qualifiers	
source		1. 1481	
rRNA		/organism="Leptospira fainei" /strain="Hurstbridge" /db_xref="taxon:48782" <1..>1481 /product="16S ribosomal RNA"	
BASE COUNT		391 a 335 c 439 g 314 t 2 others	
ORIGIN			
Query Match	96.9%; Score 1431.4; DB 1; Length 1481; Best Local Similarity 99.7%; Pred. No. 1.4e-14; Matches 1476; Conservative 0; Mismatches 1; Indels 4; Gaps 4;		
1	gatcattgctcgaactaacgctctggcggcgctctcttaacaatgtaacgtagtgg	60	
1	gatcattgctcgaactaacgctctggcggcgctctcttaacaatgtaacgtagtgg	60	
61	caatccctagcggcgaaacgggtgagttaacgct-gtlaattctcttcggagctctggata	119	
61	caatccctagcggcgaaacgggtgagttaacgct-gtlaattctcttcggagctctggata	120	
120	acttcgcgaagaagaagcctaataccgtagtcctctgtgatacacaagaattgata	179	
121	acttcgcgaagaagaagcctaataccgtagtcctctgtgatacacaagaattgata	180	
180	aagatttatctctgtgagatgtgacccggcgccgatactagctagtgtgtgagttaatgctc	239	
181	aagatttatctctgtgagatgtgacccggcgccgatactagctagtgtgtgagttaatgctc	240	
240	accaagcggaagatcgtagccggccttgagaaggtgtgccgcacaatggaactgagaca	299	
241	accaagcggaagatcgtagccggccttgagaaggtgtgccgcacaatggaactgagaca	300	
300	cggctcatactcctaagggaagcaggaagcttaagaactctgtctcaatgtaggggaaacctga	359	
301	cggctcatactcctaagggaagcaggaagcttaagaactctgtctcaatgtaggggaaacctga	360	
360	agcagcggaacggcggtgaaacgaagaagcttcggaattgtaagttcatatgaag-aga	418	
361	agcagcggaacggcggtgaaacgaagaagcttcggaattgtaagttcatatgaag-aga	420	
419	aataagcagcaactgtgatactgtacccctgaagcaccggactaacgctgccaaga	478	
421	aataagcagcaactgtgatactgtacccctgaagcaccggactaacgctgccaaga	480	
479	gccggcgtaatacgtatgtgtgcaagcgtgtgttcggaatcaattgggcgttaaggtgtcgta	538	
481	gccggcgtaatacgtatgtgtgcaagcgtgtgttcggaatcaattgggcgttaaggtgtcgta	540	
539	ggcggattgtgatacgaagctgtaaaactctggcgcacaaoccttgccctgcaacttgaac	598	
541	ggcggattgtgatacgaagctgtaaaactctggcgcacaaoccttgccctgcaacttgaac	600	
599	tacaagctcgtgaagctgtggagaaggaagtcggaattcccaagctgtagcgtgaaatcgctg	658	
601	tacaagctcgtgaagctgtggagaaggaagtcggaattcccaagctgtagcgtgaaatcgctg	660	
659	atactcgggaagcaaccacgctggcgaagggcaactctgtctgctcaaaactgaacgtggaga	718	
661	atactcgggaagcaaccacgctggcgaagggcaactctgtctgctcaaaactgaacgtggaga	720	
719	cgaaaacgttggtagttaaaacgggattagataaccgcgtaataccacgcccataaacgctgttc	778	
721	cgaaaacgttggtagttaaaacgggattagataaccgcgtaataccacgcccataaacgctgttc	780	
779	taccagctgtttgggggtttttaaacccctcagtaagaacctaaggaattgaatgaacggcct	838	
781	taccagctgtttgggggttttttaaacccctcagtaagaacctaaggaattgaatgaacggcct	840	

Qy	839	gggagcatatgctgcgaagagtgaaacccaagaagatgacgggggtccggacaagaagcg	898
Db	841	ggggacatvtagctcccaagaggaactccaaggaattgacgggggtccgacaaagcggtg	900
Qy	899	gagcatcgtgttcatcgaatgaatacccaaaaacccacccttggtgcttgaca	958
Db	901	gagcatcgtgtttaaatttcgaatgaatgatacccaaaaacccacccttggtgacatggaatc	960
Qy	959	aacatcgtagagatalabagaccttcggcgagatcgaagtcgtcga	101
Db	961	aatcatgtagagatafatmgacaccttcggcgagattcacaagctgcgcatggtgtcga	102
Qy	1019	gctcgttcgtcgaatcgtttgggttaagttcccgaaagaggaaccccatcgtatgtg	107
Db	1021	gctcgttcgtcgaatgattgggttaagttcccgcaacagcgcaaccccttatcgtatgtg	108
Qy	1079	ctaac-ttaagttcggcactgtgacgaacatcccgctgcacaacccggaagcg	113
Db	1081	ctacacatttaagttggcgactcgtatgcaaaactgcgggtgacaaacccgagagagcg	114
Qy	1138	tgaagtcacatccatcgaatgaccttatgtlccagggcgacaacgltgcataatg	119
Db	1141	tgacctcgaatccatccatmgccctttatgtccaggcgacacacgctacatg	120
Qy	1198	cagaaggttcgcgaactcgcgaagggagacatactctcaaaagtcgtcccgatc	125
Db	1201	cagaaggttcgcgaactcgcgaagggagacatactctcaaaagtcgtcccgatc	126
Qy	1258	ggggtctgcaactcgaccccaatgaagtcgaaatcgctagttaatcgggatacga	131
Db	1261	gggggtctgcaactcgaccccaatgaagtcgaaatcgctagttaatcgggatacga	132
Qy	1318	cggtgaaatagttcccggaactgtgaacacggccgtgaacacccacttgatg	137
Db	1321	cggtgaaatagttcccggaactgtgaacacggccgtgaacacccacttgatg	138
Qy	1378	cccggaagtgatcttctgttaacccgtgaaggaacagactaagtcgaactcg	143
Db	1381	cccggaagtgatcttctgttgaatccgtgaaggaacagactaagtcgaactcg	144
Qy	1438	ggtgaaagtcgtatacaagta-ccgtaaatcgaatccctcgag	1477
Db	1441	ggtgaagtcgtatcaaggtarccgtaaatccatctcttcacg	1481
RESULT	2		
LOCUS	LFA19243	1450 bp	DNA
DEFINITION	Leptospira fa1nei partial 16S rRNA gene, SSI 5402-98.		
ACCESSION	Y19243		
VERSION	Y19243.1	GI:6318184	
KEYWORDS	16S ribosomal RNA; 16S rRNA gene.		
SOURCE	Leptospira fa1nei.		
ORGANISM	Leptospira fa1nei.		
REFERENCE	Bacteria; Spirochaetales; Leptospiraceae; Leptospira.		
AUTHORS	1 (bases 1 to 1450) Petersen A.M., Krogfelt, K.A., Perolat, P., Boye, K., and Schlichting, P.		
TITLE	Leptospira fa1nei serovar Hurstbridge isolated from two Danish patients with Weil's syndrome		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 1450)		
AUTHORS	Krogfelt, K.A.		
TITLE	Direct Submission		
JOURNAL	Submitted (05-NOV-1999) K.A. Krogfelt, Statens Serum Institut, Dep. of Gastrointestinal Infections, Artillerivej 5, 2300 Copenhagen, DENMARK		
FEATURES	Location/Qualifiers		
source	1..1450		
	/organism="Leptospira fa1nei"		
	/strain="SSI 5402-98"		
	/db_xref="taxon:48782"		



	Matches	1384:	Conservative	0:	Mismatches	74:	Indels	4:	Gaps	4:
OY	1	gatcatggtcctgaagactaacgcttgccggtgcgtcttcaaca	tgcaagtcgacggtgtag	60						
Db	15	GATCTGTGGCTCGAAGAACTAACGCTGGCGCGCCTCTTAAACA	CAATGAACGGGAGTGTG	74						
OY	61	caatctcctaggcgacggaacgggtgtgtagtaaacgt	-gttaactcttcctccgagctgtgata	119						
Db	75	CAATTCCTACGCGCGGCAACCGGTGAGTAAACACTGGGTAT	CTTCCCTGTGAGTCTGGGATA	134						
OY	120	acttcgcgaagaagaaagctaataccggatagtcctgtttgtat	ctcaacaagaattgtatagtta	179						
Db	135	ACTTTCGCAAAAGGGAAGCTAATACGTAGATGGTCCCGAGAG	ATCACAGATTTTTCGGGATA	194						
OY	180	aagatttttttctgttggaatagaccgcgcgcgcgcgcgcgcgc	gc	239						
Db	195	AAGATTTTATTGCTCGGAGATGAGCCCCGCGCTCGATTAGCT	AGTTGGTGTAGTTAAGGCTC	254						
OY	240	accaagcgacgacatcgcgttagccgcgcctgtgaaggtgtgtc	tcgcgcgaacaatgaaactgagaca	299						
Db	255	ACCAAGGCGACATGCGTGGTGGCCGCTTGAGAGGGTGTTCGG	CCACAAATGGAACCTGTGAGACA	314						
OY	300	cggtccatactctctccggaagcgacgaagtaagaatctgtgcc	aatgtgggggaacccctta	359						
Db	315	CGGTCCATACCTCTTACGGGAGCAGAGTTAAAGAACTTGCTC	TAATGGGGGAGAA-CTTGA	373						
OY	360	agcagcgacgcgcgcgtgaaacggaagaagtccttcgcgtatg	taagttcatctagcca-ggaa	418						
Db	374	ACCACGACGCGCGCGGTGAACGATGAAGGCTTCGATGTAAG	AGTTCAAGTAACGAGGAA	433						
OY	419	aaataagcagcaatgtatgtatgtactgtccttaagaacccgt	ctactactgtgtgcgaaga	478						
Db	434	AAATTAAGCAGCATGTGATGATGTGTACCTGCTTAAAGCA	CCGCGCTAATACGTGCCAGCA	493						
OY	479	gcccgcgtgaatacgtatgtgtcaaggtgtgttcgcgaatcat	ctgtgcgtlaaagggtgcgtta	538						
Db	494	GCCGCGGTAAATACGTATGTGTGCAAGCGTTGTTCGGAAT	CTTGTGGGGTAAAGGTTGCGATA	553						
OY	539	ggcgcgaattgttaagtcgaagttgtgaaacctgtgcgcgcgc	ctcaaacccgttgccctgtgcattgaaac	598						
Db	554	GCGCGCATGTAAAGTCAGCTGTGTAACATCGCGGCTCACT	CCGACCCCTGCACCTTGAAC	613						
OY	599	tacaagtcctggaggttttgggaagagcgaagtgtgaaattc	ccaagttgtatgacgtgtgaaatgtgcgtag	658						
Db	614	TATGTCTCTGGAAGCTTTGGGAGAGCGCAAGTGTGAATTC	CAGGTGTAGCGGTGAATTCGCTAG	673						
OY	659	atatctgtgaaggaacacacgaatgtgcgaagcgacattgtc	gtgtctcaaaaactgacgcgtgtgagca	718						
Db	674	ATATCTGGAGGAACACACGAGCGGAGCGACATTGCTGACCT	TAACCTGACCTGTGAGCA	733						
OY	719	cgaaagcgtgtgtgtatgaacccggatagttaccgcgcgttaa	tcacagccctcaaacacgtttgttc	778						
Db	734	CGAAACCGTGGGTGAGTGAACGGGATTAGATACCCCGGTAA	TCACAGCCCTTAACAGGTTGTC	793						
OY	779	taccagttgttgggggttttlaacccttcgaagcaacctaa	gcgaatlaagtaagcgcct	838						
Db	794	TACCAATTGTTGGGGTTTTAAACCTCAGTAAAGAACCTA	AGGATTTAAATTAGACGCGCT	853						
OY	839	ggggacatgtctcgcgaagtgaaactcaaaaggaattgac	gcgggggttcgcgaacaagcgtgt	898						
Db	854	GGGAGATTAAGTCGCAAGAGTGTAAACTCAAAAGGAATT	GAGGGGGGTCCGCAACAAGCGGTG	913						
OY	899	gagacatgttgttaatttcgaatgaatacccaaaaaaacct	caccttggtgccttgaatgagatctg	958						
Db	914	GAGCAATGTGGTTTAAATTCGATGATACCCGAAAACCTC	CACCTAGCCTTGACATGGAATGG	973						
OY	959	aatcatgtatagagaataatagagccttcgcggcagaattc	caaggtgctgatatgtttgtcgtca	1018						
Db	974	AATCATGTAGAGATACATGATGACCTTGCGGCCTTCAC	AGTGTGATGTTGTGCTCA	1033						
OY	1019	gtctcgtgtcgttgagatgttgggtttaagtctccgcgaaga	gcgaagcgcaacccctatcgtatgtt	1078						
Db	1034	GCCTCGTCTGTGATGATGTTGGGTTAACTCCCGCAACG	AGCGCAACCTCTGACCTTATTTTG	1093						

QY	1079	cta -ccttaagttggcaactgtgtacgaacatgcggygtacaaaccggaaggaaggccgagga	1137
Db	1094	CCATCATTCAGTTGGGCACTGTAAAGGAACTCCGGGTACAAACCGGAGGAGCGGGGA	1153
QY	1138	tgaagcacaaccacatbqacctatgtccagaggccacacacagtgctacaaitygcgata	1197
Db	1154	TGAGCTCAATTCCTCATGCGCTTTATGCTTAGGGCAACACACAGTGTCAATATGCGCGTA	1213
QY	1198	cagaaggtctgcgaactcgtcgaagaggagatctatctctctaaagtctcggtccagttcgatt	1257
Db	1214	CAAAAGGTAGGCAATCCGAGGGGGAGCTATATCCAAAATTCGGGTCCAGTTGGATT	1273
QY	1258	ggggctcgaactcgaaccacatgaatgaatcgctagtaatcgcgagatcaagatgcg	1317
Db	1274	GGAGCTGCACATCGACATCCATGAAAGTGGAAATGCTATGATCGCGATCAGCATGCGG	1333
QY	1318	cgtgtaaatcgtctccggaacttgcacacacccgcgtcaacacacccatcgatggggagca	1377
Db	1334	CGGTAAATACGTTCCCGGACCTTGTACACACCGCCGTCACACCACTGAGTGGGAGCA	1393
QY	1378	cccggaagtggtcttctgttcaacagctaaaggagacagactaagtgtgaactcgtataagg	1437
Db	1394	CCCGAAGTGTGTTTGTGCCAACCGCAAGGAGACAGACTACTAAGTGAATCGTGAAGG	1453
QY	1438	ggtgaagtcgtatacaaggtac	1459
Db	1454	GGTGAAGTGTGTACAAAGGTACG	1475
RESULT	4		
LOCUS	L116SRDNY	1374 bp	DNA
DEFINITION	L. inadai gene for 16S ribosomal RNA (partial).		02-DEC-1993
ACCESSION	Z21634		
VERSION	Z21634.1	GI:433581	
KEYWORDS	16S ribosomal RNA.		
SOURCE	Leptospira inadai.		
ORGANISM	Bacteria; Spirochaetales; Leptospiraceae; Leptospira.		
REFERENCE	1 (bases 1 to 1374)		
AUTHORS	Hockey, J. V.		
JOURNAL	Phylogeny of Leptospiraceae and related Spirochaetes		
TITLE	Unpublished		
REFERENCE	2 (bases 1 to 1374)		
AUTHORS	Hockey, J. V.		
JOURNAL	Direct Submission		
TITLE	Submitted (09-FEB-1993) Hockey J. V., Public Health Laboratory		
JOURNAL	Service, Leptospira Reference Laboratory, Stoneham Road, Hereford,		
FEATURES	Herefordshire, United Kingdom, HRI 2BR		
SOURCE	Location/Qualifiers		
	1. 1374		
	/organism="Leptospira inadai"		
	/strain="LYME"		
	/db_xref="taxon:29506"		
RNA	<1..>1374		
	/product="16S ribosomal RNA"		
BASE COUNT	358 a 311 c 413 g 289 t	3 others	
ORIGIN			
Query Match	86.0%; Score 1270; DB 1; Length 1374;		
Best Local Similarity	96.5%; Pred. No. 3.4e-12;		
Matches 1377; Conservative	2; Mismatches 43; Indels 3; Gaps 3		
QY	16	ctaagcttgccgagcgcgtcttaacaatgaatcgagcgagggatagcaatactagcgagc	75
Db	1	STARCCTTGCGGGCGGCTTAAATCAATGCAAGTCGAGCGGGGTACCAATCACTAGCGCG	60
QY	76	aaaggttgatgaatcaagcgt-gtataatctctcggatctgggataactctcgaagaaga	134
Db	61	AACGGTGTGATTAACAGTGGGTATATCTTCTCCGAGCTTGGGATTACTTCCGAAGGAA	120

OY	135	agcctaataccggaatagtcgtgtgtggtatcccaagatttgatagtgtaagaatttatctgtc	194
Db	121	AGCTAAATACGGATACGTTACCTACTAGGTATCCAGGAATCTGATAGTAAACATTATGTGGTTG	180
OY	195	gagaatgagcccgcgccggaattgaatcagttctgggtgagtaatgctcaccagaaggacatc	254
Db	181	GAGATGAGCCCGCGGCCGATTAAGTACGTAGTTGGTGGAGGTAAAGGCTCACCAAGGCACATC	240
OY	255	ggtatgcgcgcctcgagaaggtgtccgcgcacaaatggaaactgagacacggttccatctcta	314
Db	241	GGTAGCCGGCTCAGAGGGGTGTCGGCCACCATGTGAACTGAGACACGGTCCATCTCTTA	300
OY	315	cgggaagcgcaagcttaagaatcttctgtctaaatgggggaaacccctgaagcagcagccgcg	374
Db	301	CGGAGGACACACTTAAGAACTCTTGCCTCAATGGGGGAAACCTTGAAACGACGACCGCGC	360
OY	375	tgaacgaagaagttcttcggaattgttaagttcatatagca-aggaaataagacgaacatg	433
Db	361	TGAACGAAAGAGGCTCTTCGGATTGTAAAGTTATTTAACGAGGAAAAATTAAGCAGCATG	420
OY	434	tgaatgtaactctgtccctaaagcaccgcgctaactacgttgccagcagccgcggtatatacgt	493
Db	421	TGATGATGTAACCTGCCTTAAGACACGGCTAATCTAGTCCAGACGCGCGGTATTAAGT	480
OY	494	atggttcgaagcgttgttctggaatcatltgggcgtlaaagggtgcgtgagcgttctgaagt	553
Db	481	ATGTGTGCAAGCGTTGTTCGGAAATCATTTGGCCCTAAAGGGGTGCGTAGCGGATTTGTAGT	540
OY	554	caggctgtgaaaacctgcgggtctcaaccctgcgcctcacttgaactataaagctcggagtc	613
Db	541	CAGGTGTGAANAACGCGGGCTCAACCCGAGGCGCTGACTTGAACCTACAACTCTTGAGTT	600
OY	614	tgggaaggagcaagtgtgaatctccagctgttagcgttgaaatgtcgaatactctggaggaaca	673
Db	601	TGGAGAGGCAAGTGAATTCAGAGTGTAGCGGTGAATACGTATATCTGGAGGAACA	660
OY	674	ccagatgcgaagcgagctgtctgtctccaaactgacgcctgagacgaagaacgttggtag	733
Db	661	CCAATGGCGAAGGCGAATCTTCTGACTCAAAACCTGACGCTGAGAGGCGAAGAACGTTGGTAG	720
OY	734	taacgggaattagataaccgcggttaatccacgcgccctaaagatgtgtctaccagttgtggg	793
Db	721	TAAACGGGATTTAGTATACCCCGGTAAATCCACGCCCTTAACGTTGTCTTACCACTGTGTGGG	780
OY	794	gttctaacccttcagtaacgacacctaaccgattaaagtacgcgcctctggagacatctgcgc	853
Db	781	GTTTAAACCCCTCAGTAAACGAACTTAACGGATTAAAGTAGACGCCCTCGGGACTATGTCTGC	840
OY	854	aagagctgaacatcaagaagatgtgacgggggtccgcgaacacggttgagacatgtgtttaa	913
Db	841	AAGAGTGAAACTCAAAAGGATTTGACGGGGGTCCGCACAAGCGGTGAGACATGTGGTTTAA	900
OY	914	ctcgaatataccccaanaaccttaaccttgggcttgacatggaatctgaatcatgtttagagata	973
Db	901	TTTCATATATACCGCAAAAACCTTCACTCTGGGCTTGAATGAGATCTGAATCATGTAGACATA	960
OY	974	taataagccttcgggcaaatcacaattcacaggttgtctgaatgttctgtcaagctcgttctgaga	1033
Db	961	TATATAGCCTTCGGGCAATTCACAGGTGCTGCATGGTTTCTCTCACTCGTGTCTGTGAGA	1020
OY	1034	tgtttggttlaagttccgcgaacgacgcacacctatcgtatgttcttaaccttaagttlgg	1093
Db	1021	TGTT-GGTTAAGTCCCGCACAGAGGCAACCTCACCTTATGTGTGCCATTCATTGAGTTGG	1079
OY	1094	cactggtacgaacctgtccggttgacaaacccgggggaagcggggagatggaactaaatctcta	1153
Db	1080	CACCTGTAAAGAACTGCGGTGTGCAAAACCGAGAAAGGCGGGGTGAGACGTCAAGTCTCA	1139
OY	1154	tggccttatctccaggacacacacagctgtctcaaatggcgcgtacagaggggtcgcgaact	1213
Db	1140	TGGCCTTATGTCTTAGGGCAACACAGCTGCTCAAAATGGCGGTTACAAAGGTAGCCAACT	1199
OY	1214	cgcgaaggggagtaactctctaaagltcggtcccgatlcggaattggggltctgcaactcga	1273

Db	1200	CGCGAGGGGGAGGACTATATCTCAAAAANCCGGGTCCAGTTTCGGATTGGAATCTGCACAATCGA	1259
QY	1274	cccaatgaagtcggagatcgtatgaatcggagatcagatgcgcggtgaatctccc	1333
Db	1260	CTCCATGGAAGTGGAAATGCGCTGTAATATGCGGATACGATCGCCGGGTAACTTCC	1319
QY	1334	ggacctgttacacacgcgcgcgtcacacacactgaatggggagaccggaagtgt	1388
Db	1320	GGACCTTGATACACACCGCCCTCACACCACTGATGGGGAGCACCCGAAAGTGT	1374
RESULT	5		
LOCUS	WA45165	1483 bp	DNA
DEFINITION	Leptospira weilii Cellidoni (ATCC 43285) 16S rRNA gene, partial sequence.		
ACCESSION	U12676		
VERSION	U12676.1	GI:558937	
KEYWORDS			
SOURCE	Leptospira weilii.		
ORGANISM	Bacteria: Spirochaetales; Leptospiraceae; Leptospira.		
REFERENCE	1 (bases 1 to 1480)		
AUTHORS	Rajp, D. and McClelland, M.		
TITLE	Phylogenetic evidence for horizontal transfer of an intervening sequence between species in a spirochete genus		
JOURNAL	J. Bacteriol. 176 (19), 5982-5987 (1994)		
MEDLINE	95014031		
REFERENCE	2 (bases 1 to 1483)		
AUTHORS	McClelland, M.		
TITLE	Direct Submission		
JOURNAL	Submitted (25-JUL-1994) Michael McClelland, California Institute of Biological Research, 11099 North Torrey Pines Road, La Jolla, CA 92037, USA		
FEATURES			
source	Location/Qualifiers		
	1..1483		
	/organism="Leptospira weilii"		
	/strain="serogroup Cellidoni, serovar cellidoni, strain cellidoni (ATCC 43285)"		
	/isolate="NA45"		
	/db_xref="taxon:28184"		
	<1..1483		
	/product="16S rRNA"		
BASE COUNT	382 a 331 c 452 g 315 t 3 others		
ORIGIN			
Query Match	85.6%; Score 1263.8; DB 1; Length 1483;		
Best Local Similarity	94.1%; Pred. No. 3.9e-12;		
Matches 1376; Conservative	0; Mismatches 80; Indels 6; Gaps 6;		
QY	1	gatacgtcgaagaaagcgtcgtgcgcgcgtcttaacatgcaagtcagcgagtag	60
Db	10	GATCTGGGTCAAGACTAACGCTGGCGGGCTCTTAAACATGCAAGTCACGAGTAG	69
QY	61	caatacctgcggcgaaacgggtgagtaaacagt-ggtaatctctccgagtcgagta	119
Db	70	CATATCTACGCGCGACGCGGATGATACACGTGGTAACTTCTCCGAGCTGGATA	129
QY	120	acttccgaaagaaagcctaataccggaatagtcctgttgatcacaaagattgata	179
Db	130	ACTTTCCGAAGAGGGAGCTAATATGAGATGTCGCGAGAGTCATATGATTTTCCGGTA	189
QY	180	aagatttatctgcttgagatgagcccgcgccgattagctatggtgaggtgaatg	239
Db	190	AAGATTATTTGTCGAGAGCTGAGCCCGCCCGCATTAAGCTAATTGGTGAAGTAAGCTC	249
QY	240	accaagcgagcgcgtgtagccgcgcctggaagaggtgctgcgcacaaatgaa	299
Db	250	ACCAAGGCGACATCGTAGCGCGCTGAGAGGGGTGTGGCCACAATGGAATGAGACA	309
QY	300	cggtccatactctcaccggagagcagcagttaaatctgtctcaatlggggaaacctga	359

Db	310	CGGTCCATACCT -CTACGGGAGGACAGATTAAAGATCTTGCTCAATGGGGGAACCCCTGA	368
Oy	360	agcaagcagcgcgcgtgaaagaaaggtcttcggatctgttaagttaataagca -gga	418
Db	369	AGCAGGACCCCGCGAAGCATGAAGGCTTCGGATTGTAAATTCATAATACGAGGAA	428
Oy	419	aaataagcaagcaatgtgatgatcgttaccccgctaaagcaccggttaatactatgcagca	478
Db	429	AAATTAACACGACATGTGATGATGATACCTCGCTAAAGACGCGGTAACTAGTGCACGA	488
Oy	479	gccgcgtaatacgtatgtagcgaagcgttgttcggaatcalttggcgtgaaaggtgcgt	538
Db	489	CGCGGGTAATACGTATGTTGTCGAAGCGTGTTCGGAAATCATTTGGCGCTTAAAGGTGTGA	548
Oy	539	ggcggaattgttaagtcaggttgtgaaacatcgcggtgtcaaccgctgcgcgcgaacttgaac	598
Db	549	GGCGGACATTAAGTCAGGGGTGAACACGCGGCTCACTGCGACGCTGCATTGAAAC	608
Oy	599	tacaagctctgaagctctggagagagcgaatggaattccagagtgtaacggtgaaatgcgtag	658
Db	609	TATGTGCTCGAATTTGGGAGGCAAGTGAATTCGAGGTGAGCGGTGAATGCGGTAG	668
Oy	659	atalcgtgaggaacaacccatgtgcgaagcgcaactgtcgtgtccaaactgcaagcgtgaaaga	718
Db	669	ATAATCTGGAAGAACACCAAGTGGGGAAGGGGACTTCTGTGACCTAAACATGACGCTGAGCA	728
Oy	719	cgaagcgttggttagtaaaacggaataga taaccggtaatccacgcgcttaaacgcttgc	778
Db	729	CGAAGCGTGGATGTGACGAGGATTAAGTACC CGGTATATCCAGCCCTTAACGTTGTC	788
Oy	779	taccagttgttgggggttlttaaccctcag taacgaactaacyga ttaagttagacgcct	838
Db	789	TACAGTGTGGGGGTTTAACTTAACTCACTAATACGA -CTAACGGATTAAAGACGCCCT	847
Oy	839	ggggacatctcgcgaagatgaaactcaagaatgtgacgggggttccgcgaagcggtg	898
Db	848	GGGACATATGCTCGCAGAAGTGAACACTCAAGGAATTAAGGGGGTNNNCAAGCGGTG	907
Oy	899	gagcagtgtgttaaatcga tga laacccaanaaacctcaactcgtggcttgacatgatactg	958
Db	908	GAGCATGTGGTTTAATTCGATGATACGCGAAAACCTCACTAGCGCTTACGATGGAGTGG	967
Oy	959	aatcatgtagagataatagtccttcgggcagatcaacaggtgcacatgttgttctgcga	1018
Db	968	AATATGTAGAGATCACTVGA -CTTCGGGCGGCTTACAGGCTGCAATGGTGTGCTCA	1026
Oy	1019	gctcgtctgcgaaatgtcttggttgaatcccgcaacgaagcgaaacccctatcgtatgtgtg	1078
Db	1027	GCTCGTGCATGAGATGATGTGGGTTAAGTCCCGCAACGACGCAACCCATCCTTATGTTG	1086
Oy	1079	cta -ccttaagcttggaactcgtgtacgaaaactcgcgtgtgacaaaccggggaggaagcgagg	1137
Db	1087	CCATCATTTACTTGGGCACTGTAAAGAACTGCGGGTGAACAAACCGGAGGAAGGCGGGGA	1146
Oy	1138	tgaagtcacaaatcccaatgagcgccttatgtccagggcacaacacggtgtlaaactgagccgta	1197
Db	1147	TGACGTCAAATCCTCATGCGCTTTATGTCTTAAGGCAACACAGTGTTACAAATGGCGGTGA	1206
Oy	1198	caggaggttcgcaaacctcgcaagaggaagcctaactcctaaagt cgtgtccagcttcggaat	1257
Db	1207	CAAAAGGTAAAGCAACTCGCGAGGGGAGCTTAATCTCAAAAAGCGGCTCCAGTTGGGATT	1266
Oy	1258	ggggctctcaaatcgaaccccaatgaagtcggaatcgctagttaactcgggtgacgtgcga	1317
Db	1267	GGAGTCTGACATGCACTCCATGTAAGTCGGAATTCGTAATATCGGGATACGACATGCCG	1326
Oy	1318	cggtaataacgtctcccgagcctgttaacacacgcgcgctcacacacactgagtggggagca	1377
Db	1327	CGGTAAATACGTTCCCGGACCTTATACACAACGCGCGTCAACACCTAGTGGGGAGCA	1386
Oy	1378	ccggaagtgttcttgttaaccgttaaggagacagactactaaagtgaaaactcgttaaggg	1437

Db	1387	CCCCAGACGGGTCTTTGCCAACCCGTAAGAAGACAGACTACTAAAGTGAAACATCGTGAACGG	1446
QY	1438	ggatgaagtcgataacaaggtacc	1459
Db	1447	GGTGAAGTCGTAAACAGGTAGC	1468
RESULT	6		
LOCUS	WA5216S	1489 bp	DNA
DEFINITION	Leptospira weilii Worsfold 16S rRNA gene, partial sequence.		
ACCESSION	U12677		
VERSION	U12677.1	GI:558938	
KEYWORDS			
SOURCE			
ORGANISM	Leptospira weilii.		
REFERENCE	Bacteria; Spirochaetales; Leptospiraceae; Leptospira.		
AUTHORS	1 (bases 1 to 1480)		
TITLE	Ralph.D. and McClelland.M.		
JOURNAL	Phylogenetic evidence for horizontal transfer of an intervening		
REFERENCE	sequence between species in a spirochete genus		
AUTHORS	J. Bacteriol. (1994) In press		
TITLE	2 (bases 1 to 1489)		
JOURNAL	McClelland.M.		
REFERENCE	Submitted Submission		
AUTHORS	Directed (25-JUL-1994) Michael McClelland, California Institute		
TITLE	Biological Research, 11099 North Torrey Pines Road, La Jolla, CA		
JOURNAL	92037, USA		
FEATURES			
source	Location/Qualifiers		
	1..1489		
	/organism="Leptospira weilii"		
	/strain="serogroup Hebdomadis, subserogroup Borlincana,		
	serovar worfoldi, strain Worsfold"		
	/isolate="WA52"		
	/db_xref="taxon:28184"		
	<1..1489		
	/product="16S rRNA"		
BASE COUNT	383 a 332 c 455 g 316 t		3 others
ORIGIN			
Query Match	85.6%; Score 1263.8; DB 1: Length 1489;		
Best Local Similarity	94.1%; Pred. No. 3.9e-12;		
Matches 1376; Conservative	0; Mismatches 80; Indels 6; Gaps 6		
QY	1	gatacaggtcagaactaacgcctgctggcgagcgcttaacaatcagcaagtcgagcg99gtag	60
Db	16	GATCCCTGGCTCAGAACTAACGCTGGGGGGCGCTTAAACATGCAAGTAAAGCGGAGT	75
QY	61	caatactacgcgagcgaggtgagtaaacgct-ggtlaactctctccgagtcgagata	119
Db	76	CATACTCAAGCGCGGAGGAGGCGGTGAGTAAACGCGGGGTAACTCTCCGAGTCTGGGATA	135
QY	120	acttccgaaaggaagaaactaatacgcgataagtcctcgttgatcacagaatttgatagta	179
Db	136	ACTTTCGGAAGGAGGAGCTAATCTGATGCTCCGAGAGGTCATATGATTTTCGGGTA	195
QY	180	aagattatgctcttgagatagagcccgcgccgattagctagcttggtgagtaatagctc	239
Db	196	AAGATTATTGCTCGAGAGCTGAGCCCGCCGCCATTAGCTAGTGGTGAGAGTAATGGCT	255
QY	240	accaagcgagacgatacggtagccggtcgtgagaggtgltccggccacaatggaaactagaca	299
Db	256	ACCAAGCGACGATCGGTACCGGCTCGAAGAGGGTGTTCGGCCACATGGAATGTGAGACA	315
QY	300	cggtccatactcttaaggggggagcagtgtaagaatctgctcaatgggggaaacctga	359
Db	316	CGGTCCATTAAT-CTACGGAGGAGCAGCTTAAGAAATCTCTCAATGGGGGAACCTGA	374
QY	360	agagcagagcgcggtbaacgaaagagcttcggaattgaaagttaactaagga-ca-gga	418
Db	375	AGCAGCAGACCCCGGTGAGATGAAAGGCTTTCGATGTTGAAAGTTCAATTAACAGAGGA	434



OY	419	aataaagcaacaaatgataatgtatccctgacccgccaacaaagcccgccaaacttaactgctgcacga	478
Db	435	AAATTAAGACGAGATGTAATATGTACTCGCTAAAGCAACCGGCTACTACTGCTGCACGA	494
OY	479	gcccgggtaataactgatactgacgaagcgtgtcttcgaaatcaattggggcgtlaaagggtgcga	538
Db	495	GCCCGGGAATATAGGTATGGTGCAAAGCGTTGTTCGAAATCATTTGGGCGTAAAGGTCGCTA	554
OY	539	ggcggattctgtaagtcaggctgtgtgaaaaactgcccggctcaaccggctgacctgtgaac	598
Db	555	GGCCGACATGTAAAGTCAGGCGTGAATAACGCGGGCTCAACTGCGACGCGTGCCTTGAAC	614
OY	599	tacaagcctcgagatctggggggaaggaagatgtccaaagttaacggttgaaatgtcag	658
Db	615	TATGTCTGTGGAGTTTGGGAGAGCAAGTGGAAATTCAGAGTTACGGGTGAATATCGTAG	674
OY	659	atactctggaggaacaacgaatgagcgaagcgcaactctgtctctcaaaaactgtaacgtlaagca	718
Db	675	ATATCTGGAGGAACACCAAGTGGGGAAGGGGACATTGCTGGCTTAATAACTGACGCTGAGCA	734
OY	719	cgaaagcgtgtggttagtaaaacggggaattagataccccggtatccaaagccctaaagctgtc	778
Db	735	CGAAAGCGTGGTATAGTCAAGCGGATTAAGTATCCCGCTAATCCAGCCCTTAACCTTTC	794
OY	779	tacaagctgtctgggggtcttaaacccctcaatlaacgaaccctaaacggaattlaagtgaacgct	838
Db	795	TACACGTTGTTGGGGGTTTAAACCTTCAGTAACGA-CTAACGGATTAAGTGACCGCCT	853
OY	839	ggggactaagctcgcgaagatgaaactcaaaaggaattgacgggggtlccgcacaagcgtg	898
Db	854	GGGACATAGCTCCCAAGACTGAACCTCAAGGAATTTGACGGGGTNNNCACAAACCGGTG	913
OY	899	gagatctgtcttaacttcgatactgataccccaanaactcaaccggcctgaactgtatctg	958
Db	914	GAGCATGTGTTTAATTTCATGTATACGCCGAANAACCTCACTAAGGCTTGACATGGAGTGG	973
OY	959	aatactgtagagataatactgaccccttcgggcgaatltcacaggtgtctgaactgtgtctgcga	1018
Db	974	AATATATAGAGATACATGAG-CTTCGGGGCGCTTCACAGTGTGATGATGTTCTGTCGA	1032
OY	1019	gctctgtctgtgagatgtgtgtgtlaagctcccgaaagacgaaccctatctgtatctgt	1078
Db	1033	GCTCGTGTGCGAGATTTTGGTTAAGTCCCGCAACGACCGCAACCCCTCACTTATGTTG	1092
OY	1079	cta-ccctlaagttctggcaactggtacgaaacctgcggctgacaaaccgggaagcgcgga	1137
Db	1093	CCATCATTTAAGTTGGGCACTCGTAAGGAACCTGCCGTTGACAAACCGAGGAAGCGGGGA	1152
OY	1138	tgaactcaaatctcatalgtgccccttaactgtccagggccacacggtctacaatalgtgcga	1197
Db	1153	TGACGCTCAAAATCCCATGAGGCTTTATGTCTAGGGCAACACAGTGTCTCAAAAGGCGCGTA	1212
OY	1198	cagagggctgcacaactcgcgaagagggagacataactctaaagcttggtccagttctgagatt	1257
Db	1213	CAAAAGGTAGCCAACTGCGCAGGGGGAGCTAAATCTCAAAAACCGGTCGCCAGTTGGGATT	1272
OY	1258	ggggctctgaactcgcaccctgaagatcggaaatcgctagtaatactgcggaactgaacgtgcg	1317
Db	1273	GGAGTCTGCAAACTGCATCCATGAATCGGAATCGCTAGTATCCCGATCAGCATGCGG	1332
OY	1318	cggatgaatacgtctcccggaactctgtacaacccggcgtcaaacacactgaagtggggaaga	1377
Db	1333	CGGTGAATACGTTTCCCGACCTGTGTACACGCGCCGTCACACCACTGAAGTGGGAGCA	1392
OY	1378	ccggaagctgtctctgtctgaacgctlaagggacgaactaactaaggttgaactctglaaagg	1437
Db	1393	CCCGAAGTGGCTTTGGCAACCGTAAGGAAGACAGACTACTAAGGTGAACCTGTGAAGG	1452
OY	1438	ggtgaagtctgaacaaagctac	1459
Db	1453	GGTGAAGTCGTAAACGAAGTATG	1474

[illegible]

```

OY 539 ggcggatttgtaagtcagctgctgaaacatcgccggtcacaacccgtggtcctcacttgaaac 558
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 579 GGGGGAATGTAATCTCAGGTGCTGAAAATCTCGGGCTCACTCCAGCGCTCAGCTGAAAC 638
OY 599 tacaagctcggagtttggaagagcgaaatggaattccaggtgtgaagcgtgaaatcgtag 658
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 639 TATGTGCTGGAGTGTGGAGAGCAAGTGAATTCAGGTGTAGCGGTCAAAATGCGTAG 698
OY 659 atactctggaagaacacacagctggcgaagcgactgtcgtctcaaacctgaacgtcga 718
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 699 ATATCTGGAGAAACACCACTGCGCAAGCCACTTGCTGGCTTAATACTGAAGCTGAGACA 758
OY 719 cgaagcgtggtgtagtaaacggaatagatacccggttaatcaacgcttaaacgtgtgc 778
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 759 CGAAGCGCTGGTGTAGTGAACGGGATTTAGATACCCGGTAAATCCAGCCCTAAAGCTTGTC 818
OY 779 taacagttgttggggttttaacctcaagtaacgaacactgaacgattagaacgcct 838
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 819 TACCAAGTTGGGGGTTTAAACCTCAGTAAACGAH-CTAAACGATTAAGTAGACCCCT 877
OY 839 ggggactatgctcgcaagagtgaaactcaagaagattgaacgggggtccgcaaacggctg 898
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 878 GGGGACTATGCTCCGCAAGAGTGAACCTCAAGGAATTGACGGGGTNNMCACAAGCGGTG 937
OY 899 gagcattgtttaaattcctatgatacccaaaaactcaactcgtggttgacatgactctg 958
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 938 GAGCATGTGGTTTAAATGTGATGATACGGCAAAACCTCAGCTGAGCTTGACATGAGATGG 997
OY 959 aatcatgtagaagataatgaacccctcggcgcaatcacaagtgctgcaatggtctgcata 1018
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 998 AATCATGTGATGATACATGAG-CTTCGGGGCGCTTCACAGGTGCTGATGGTGTGCTGCA 1056
OY 1019 gctcgtcgtgtagaagtggtggttaagtcaccgcaacggcgcaacccctatcgtatgtg 1078
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1057 GCTCGTGTGATGAGATGTTGGGTTAAGTCCCGCAACGAGCCGACCCACCTTATGTTG 1116
OY 1079 cta-cccttaagttggaactgtgacgaacatcgctgacaaacgggaagggcgcgga 1137
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1117 CCATCATTTAGTTGGGCACTCTTAAGGAATCTCGGTGACAAACCGGAGGAAGCGGGGA 1176
OY 1138 tgaagtcacaactccatagccttatatgtccaggcgcaacacgctgctacatgctcgata 1197
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1177 TGACGTAATATCTCTATGCGCTTATGTAGGCAACACAGTCTTACAAAGCGCGGTA 1236
OY 1198 cagagtggtcgaacactcgaagaaggagagtaactccttaaaagtcgtccagttcgatc 1257
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1237 CAAGAAGTAGCCAACTCGCAGAGGGAGCTAATCTCAAAAAGCCGCTCCAGTTGGAAT 1296
OY 1258 ggggtctgaacatcgaccccatgaagtcggaatcgtagtaatcgcgatcagcatgccc 1317
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1297 GGGGTCTGCACTCGACTCATGAAGTCCGAATCGCTAGTAATCGGGATCGAGCATGCCG 1356
OY 1318 cggltgaatacgttcccggaactgttacacacgcccgttacacacactgaagtgggagga 1377
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1357 CGGTGATATAGTTCCCGGACCTTGTACACACCCCGCTCACACACCTTAAGTGGAGGA 1416
OY 1378 ccggaagtggtcttggttaacgtaaggaagacagactactaaagtgaactcgtlaaagg 1437
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1417 CCGGAAGTGGTCTTTGGCAACCCCAAGGAAGCAGACTACTAAGGTAAACTCTGTGAAGG 1476
OY 1438 ggtgaagctgtaacaagtlac 1459
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1477 GGTGAAGTCTTAACAAGTAGC 1498

```

```

RESULT 8
LB012669 1489 bp DNA BCT 18-FEB-1995
LOCUS LB012669 borgeteterseni11627 Burgos 16S rRNA gene, partial
DEFINITION sequence.
ACCESSION U12669
VERSION U12669.1 GI:558928

```

```

KEYWORDS
SOURCE Leptospira borgeteterseni11.
ORGANISM Leptospira borgeteterseni11.
Bacteria; Spirochaetales; Leptospiraceae; Leptospira.
REFERENCE
AUTHORS 1 (bases 1 to 1489)
TITLE Ralph.D. and McClelland,M.
JOURNAL Phylogenetic evidence for horizontal transfer of an intervening
          sequence between species in a spirochete genus
          J. Bacteriol. 176 (19), 5982-5987 (1994)
          95014031
MEDLINE 2 (bases 1 to 1489)
REFERENCE McClelland,M.
AUTHORS Direct Submission
TITLE Submitted (25-JUL-1994) Michael McClelland, California Institute of
          Biological Research, 11099 North Torrey Pines Road, La Jolla, CA
          92037, USA
FEATURES
          source
          1..1489
            /organism="Leptospira borgeteterseni11"
            /strain="serovar balcanica, strain 1627 Burgos"
            /isolate="BD30"
            /db_xref="taxon:174"
            <1..1489
            /product="16S rRNA"
BASE COUNT 385 a 330 c 454 g 311 t 9 others
ORIGIN
Query Match 85.4%; Score 1261; DB 1; Length 1489;
Best Local Similarity 93.8%; Pred. No. 4,3e-12;
Matches 1372; Conservative 0; Mismatches 84; Indels 6; Gaps 6;
OY 1 gatcatgctcagaactaacgcttgcgcgcgtcttaaacatgcaagtcgagcgggtgag 60
Db 16 GATCTGCGTCAGAACTAACGCTGCGCGCCGCTTAAACATCAACATCAAGCGGAGTAG 75
OY 61 caataactagcggcggaacgggtgtagtaacagt-ggtaactctcccgagctcggagata 119
Db 76 CAATATCTCAGCGGCAACGGGTGATACACAGTGGTAACTTCTCCGAGCTCGGATGA 135
OY 120 acttcgaaagaaagcctaataccggaatagctcgttgcatacgaagattgatalgta 179
Db 136 ACTTTCCGAAGGGGAGACTAATFACTGTGATGCTCCGAGAGGTCTATGAGATTTTCCGGGTA 195
OY 180 aagattatctcttgagagatgaagcccgcgcgatagctagttggtgaggtaaatggtc 239
Db 196 AAGATTTATTGCTCGAGATGAGCCCGCGCGGATTAAGCTAAGTGTGAGGTAATGGCTC 255
OY 240 accaagcgacgaatcgtagtcgcgcgcctgagaggtgtccgcgcacaaatggaactgagaca 299
Db 256 ACCAAGCGGAGATCGTATGCCCGGCTGAGAGGGTTCGGGCACAAATGGAATGAGACA 315
OY 300 cggltcatalactcctacggaagcgacgaltgaagaatctctcctcaatgggggaaacctgta 359
Db 316 CGGTCTATACT-CTACGGGAGGAGCAGCACTTAAGATCTTGTCTCAATGGGGGAACCTTGA 374
OY 360 agcagcgagcgcggtgaaagaaggtcttcgattgtaaatcttaagga-9gaa 418
Db 375 AGCAGCGAGCGCGCTGACGATGAAGGTCTTGGATTTGAAGTTCAATTAAGCAGGAA 434
OY 419 aataagcgcaatgttgattgattgtaactgacaaagcaacggcgttaactaactgtccaga 478
Db 435 AATATAGACAGCATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 494
OY 479 gccgcgtaatacgtatgtagcagcgtgtgtcgaatcatatggcgttaaaaggtggtgta 538
Db 495 GCGCGGTAAATAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 554
OY 539 ggcggatttgtagtaagtgtagaaacatcgcggttcacacccgtgacctgcaacttgaac 598
Db 555 GCGGACATGTAAAGTCAAGGTGAAGAACTCGCGGCTCAACTGCAAGCTGCACTTGAAC 614
OY 599 tacaagctcggagtttggaagagcgcaagtggaattccaggtgtgtagcgtgtaaaatgcgtag 658

```



QY	Db	1318	1313	1303	1293	1283	1273	1263	1253	1243	1233	1223	1213	1203	1193	1183	1173	1163	1153	1143	1133	1123	1113	1103	1093	1083	1073	1063	1053	1043	1033	1023	1013	1003	993	983	973	963	953	943	933	923	913	903	893	883	873	863	853	843	833	823	813	803	793	783	773	763	753	743	733	723	713	703	693	683	673	663	653	643	633	623	613	603	593	583	573	563	553	543	533	523	513	503	493	483	473	463	453	443	433	423	413	403	393	383	373	363	353	343	333	323	313	303	293	283	273	263	253	243	233	223	213	203	193	183	173	163	153	143	133	123	113	103	93	83	73	63	53	43	33	23	13	3
QY	Db	1318	1313	1303	1293	1283	1273	1263	1253	1243	1233	1223	1213	1203	1193	1183	1173	1163	1153	1143	1133	1123	1113	1103	1093	1083	1073	1063	1053	1043	1033	1023	1013	1003	993	983	973	963	953	943	933	923	913	903	893	883	873	863	853	843	833	823	813	803	793	783	773	763	753	743	733	723	713	703	693	683	673	663	653	643	633	623	613	603	593	583	573	563	553	543	533	523	513	503	493	483	473	463	453	443	433	423	413	403	393	383	373	363	353	343	333	323	313	303	293	283	273	263	253	243	233	223	213	203	193	183	173	163	153	143	133	123	113	103	93	83	73	63	53	43	33	23	13	3
QY	Db	1318	1313	1303	1293	1283	1273	1263	1253	1243	1233	1223	1213	1203	1193	1183	1173	1163	1153	1143	1133	1123	1113	1103	1093	1083	1073	1063	1053	1043	1033	1023	1013	1003	993	983	973	963	953	943	933	923	913	903	893	883	873	863	853	843	833	823	813	803	793	783	773	763	753	743	733	723	713	703	693	683	673	663	653	643	633	623	613	603	593	583	573	563	553	543	533	523	513	503	493	483	473	463	453	443	433	423	413	403	393	383	373	363	353	343	333	323	313	303	293	283	273	263	253	243	233	223	213	203	193	183	173	163	153	143	133	123	113	103	93	83	73	63	53	43	33	23	13	3
QY	Db	1318	1313	1303	1293	1283	1273	1263	1253	1243	1233	1223	1213	1203	1193	1183	1173	1163	1153	1143	1133	1123	1113	1103	1093	1083	1073	1063	1053	1043	1033	1023	1013	1003	993	983	973	963	953	943	933	923	913	903	893	883	873	863	853	843	833	823	813	803	793	783	773	763	753	743	733	723	713	703	693	683	673	663	653	643	633	623	613	603	593	583	573	563	553	543	533	523	513	503	493	483	473	463	453	443	43																																											

JOURNAL	sequence	species	in a	splirochete	genus
MEDLINE	9: Bacteriol.	176 (19),	5982-5987	(1994)	
REFERENCE	95014031				
AUTHORS	2 (bases 1 to 1516)				
TITLE	McClelland, M.				
JOURNAL	Submitted (25-JUL-1994)	Michael	McClelland,	California	Institute of
	Biological	Research,	11099 North	Torrey Pines	Road, La Jolla, CA
	92037, USA				
FEATURES	Location/Qualifiers				
source	1..1516				
	/organism="Leptospira noguchii"				
	/strain="serogroup Autummalis, subserogroup Fort Bragg, serovar fortbragg, strain Fort Bragg"				
	/isolate="NB36"				
	/db_xref="taxon:28182"				
rRNA	<1..1516				
	/product="16S rRNA"				
BASE COUNT	392 a	342 c	462 g	314 t	6 others
ORIGIN					
Query Match	85.4%	Score 1260.8;	DB 1:	Length 1516;	
Best Local Similarity	93.9%;	Pred. No. 4.2e-12;			
Matches 1373:	Conservative	0: Mismatches	83;	Indels	6; Gaps
1	gatactgctcgaagactaacgccttgccgagcgctcttaaacatgcgaagtcgagcgagtag	60			
43	gATCTGGCTCAGAACTAACGCTGGCGCGCGCTTTAAACATGCAAGTAAAGCGAGTAG	102			
61	caataccttagcggcggaacggggtgagtaaacagt-gtaactcttcctccagctcgggata	119			
103	CAATCTCAGCGCGCAACGGGAGTAACTACACCTGGGTATCTTCCNCGAGTCTGGGATA	162			
120	acttcgcaaaagaaagcgaataccggaatagccctgcttgatcaacaagttgataagata	179			
163	ACTTTCCGAAAGGAGAGTAATCTAGTAGTCCGAGAGATCATAGATTTTTCGGGATA	222			
180	aagatttatctgcttgagagatgagcccgcgcgacatgaagtagcttggtgagtaatgagct	239			
223	AAGATTCTATGTGCTCGAGATGAGCCCGCGCTCGATTAGCTAGTGGTAAAGGCTT	282			
240	accaaagcgagagatctggttagcccgccctggaagaggtgtgccgccaacaatgaaactgaa	299			
283	ACCAAGGGGACGATGGTAGCCGGCTGTGAGAGGGTGTTCGGCCACATGGAACCTGAGACA	342			
300	cggtcatcactcctcactcagggagcgacagcgttaagaaatctgtccaaatgggggaaacctga	359			
343	CGGTCCATACT-CTACGGGAGGCGACAGTTAAGATCTTGTCTCAATGGGGGGAACCTCTGA	401			
360	agcagcgagcgccgctggaacggaagaagctctcggatttgtaagtcatatgaaca-ggaa	418			
402	AGCACGACGCGCGCTGAAACGATGAAGGCTTCGGAATTTAAAGTTCAATTAAGCAGAGGAA	461			
419	aaataagaagcaatgtgatatgtatgctacccgtcctaagaacccgcttaactagtgccagta	478			
462	AAATTAAGCAGCATGTGATGATGATGTACCTGCTTAAGCACCGGCTTACTACGCGCAGCA	521			
479	gcccgagtaatacgtatggttgcaagcgttcttcggaatacatctgggctgtaaaaggttcgta	538			
522	GCCCGGTAATACGATGGTGCAGAGGTTGTTCCGAATCATTTGGGGGTAAAGGATGCGTA	581			
539	ggcggaatttgtaagtcaaggtgtgaaaaactcgcggtctcaaacccgttgccctgcacttgaac	598			
582	GGCGGACATGTGAAGTCAGGTTGAAAAAATCGGGGCTCAACTCCACACCGACCTTGAAC	641			
599	tacaagctcgaagcttggtggaagagcgaaagtggaaatcccaagcttgtaggtgaaatgcgtag	658			
642	TATGTGCTCTGAGTTTGGAGAGGCAAGTGAATTCACAGTGTAGCGGTGAATATGCGTAG	701			
659	atactctggaagaaaccagctgagcgaagcgactgtcgtgctcaaaaactgaacgtcgaagca	718			
702	ATATGTGAGGAACACCAATGTGGCAAGGCACTTGCTGCTTAAACATCACCTGTGAGGCA	761			

QY 719 cgaagcggtggaagtaaacgagattagatacccggttaaccgagccctaaacgtgtc 778  
 Db 762 CGAAAGCGGTGGAGTGAACGGGATAGATTGCCCGGCTTAATCCAGCGCCCTTAACGGTTGTG 821  
 QY 779 taccagctgttgggggttttaaccctcaatgaacgaacctaacgattagtagccgc 838  
 Db 822 TACCAAGTTGGGGGTTTAAACCTTACCTAGTACGAA-CTAACGGATTAACTAGACGCCCT 880  
 QY 839 ggggaactatctcgcaagaatgaactcaagaagaaatgaaggggtccgcaagcggtg 898  
 Db 881 GGGGACATATGCTGCCAAGAGTAACAAAGAACTGAAGGAGGCGGCTNNNCAACAAGCGGTG 940  
 QY 899 gaggatgtgttaattcgtatgatacccaaaaacccctcactgggtctgacatgacgtc 958  
 Db 941 GAGCATGTGGTTTAAATTCGATGATACCGAAACCTCACCCTAGGCTTGATGATGAGTGG 1000  
 QY 959 aatacgtagaagataatgaaccttcgggacagatcaaggtgtcgtatgtgtcgtca 1018  
 Db 1001 AATCATGTAGAGATACATGAG-CTTCGGGGCGCTTCACAGGTGCTGCATGGTTGCTGCA 1059  
 QY 1019 gctcgtcgtgagatgttgggttaagtcgccgcaagcgagcgcaaccctatcgtatgtg 1078  
 Db 1060 GCTGCTGTCGAGATGTTGGGTTAAGTCCGCAACGAGCGCAACCTCACTTATGTNG 1119  
 QY 1079 cta-ccctaaagtggacatgtgtacgaactgcccgtgacaaacggagggagggcgagga 1137  
 Db 1120 CCATCATTCAGTTGGGCTGCTGTAGAGAACTGCGGTACAAACCGAGAGAGCGGGGA 1179  
 QY 1138 tgaactcaaacctccatagcttctatgtccagggccacacacacgtgtcctaaatggccgata 1197  
 Db 1180 TGAGCTCAATTCCTCATGAGGCTTTATGTCTAGGGCAACACACGTCGTCAATGGCGGGTA 1239  
 QY 1198 caagaggttcgccaactcgcaagggaggaactaactcctaaagtgtcgtccagttcgatt 1257  
 Db 1240 CAAGGGTAGGCCAACTCCGAGGGGAGCTAATCTCAAAAAGCGCGTCCAGTTGGGAT 1299  
 QY 1258 ggggtcgaactcgaccctcgaactgaagtgcgaatcgtaactcgagatcagcatcg 1317  
 Db 1300 GAGCTGTCGAACCTGCATCCATGAGTGGAAATGCGTAAATGCGGATCAGCATGCGG 1359  
 QY 1318 cgggtgaatacgttcccggaacctgtacacacggccgttcaacacacactgagtgggagca 1377  
 Db 1360 CGGTAAATACGTTCCCGGACCTTTGTACACACCGCCGTACACACCTGAGAGTGGGAGCA 1419  
 QY 1378 cccgaaggtggtctgttgaactgaagagagacagactaactgaagtgaactcgtaaagg 1437  
 Db 1420 CCCGAGGTGGTCTTTGCCAACCGCAAGGAGCAGACTACTAAGTGAACCTGTAAGGG 1479  
 QY 1438 ggtgaagtcgtacaaggtacc 1459  
 Db 1480 GGTTGAAGTGTGAACAAGTATG 1501  
 RESULT 10  
 LB16SRDNP 1425 bp DNA 02-DEC-1993  
 LOCUS L.borpetersenii gene for 16S ribosomal RNA (partial).  
 DEFINITION 221630.1 GI:433579  
 ACCESSION 221630.1 GI:433579  
 VERSION 16S ribosomal RNA.  
 KEYWORDS Leptospira borgpetersenii.  
 SOURCE Leptospira borgpetersenii.  
 ORGANISM Bacteria; Spirochaetales; Leptospiraceae; Leptospira.  
 REFERENCE 1 (bases 1 to 1425)  
 AUTHORS Hookey J.V.  
 TITLE Phylogeny of Leptospiraceae and related Spirochaetes  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 1425)  
 AUTHORS Hookey J.V.  
 TITLE Direct Submission  
 JOURNAL Submitted (09-FEB-1993) Hookey J.V., Public Health Laboratory  
 Service, Leptospira Reference Laboratory, Stonebow Road, Hereford,

Herefordshire, United Kingdom, HRI 2BR  
 Location/Qualifiers  
 source 1.1425  
 /organism="Leptospira borgpetersenii"  
 /strain="JAVANICA VELDRAI BATAVIA 46"  
 /db\_xref="taxon:174"  
 <1..>1425  
 /product="16S ribosomal RNA"  
 BASE COUNT 369 a 327 c 430 g 297 t 2 others  
 ORIGIN  
 Query Match 84.9%; Score 1254.6; DB 1; Length 1425;  
 Best Local Similarity 94.2%; Pred. No. 5,6e-12;  
 Matches 1343; Conservative 2; Mismatches 76; Indels 4; Gaps 4;  
 QY 16 ctaacgttggcggtcgttcttaacatgcaagtc-gagcggtgttaacatcctagcggc 74  
 Db 1 STARCGCTGGCGGCGCTTTAAACATCCAAATCCAAAGCGGAGTACCAATCTACAGGCG 60  
 QY 75 gaacgggtgagtaacagct-gtlaatcttcctccgaagtcgtgataacttcgaaagga 133  
 Db 61 GAACGGGTGAGTAAACAGTGGGTAATCTCTCGAGTCTGGGATTAATCTCCGAAGAGG 120  
 QY 134 aagctaaaccgagtagtctcgtgtgatacaagaattgtaggtlaaagattatgtc 193  
 Db 121 GAGCTAATACGTGATAGTCCCGAGAGGTGATAGATTCTTTCGGGTAAAGATTATGCTC 180  
 QY 194 gggatagaacggcggtgaggttagctagttgtgaggtlaaggtcacaaggcgagat 253  
 Db 181 GGAATAGACCCCGCGCCGCGATTACCTGTTGGTGAAGTAATGCTCTCCAAAGCGAGCAT 240  
 QY 254 cggtagccggtctgagaggtgttcggccacaaatgaaactgagaacagcgtcactcct 313  
 Db 241 CGGTAGCGCGGCTGAGAGGGGTTCGGGCCCAATGAACTGAGACAGGTCATACCTCT 300  
 QY 314 acgggaggtcagcaatlaagaactctgtcctcaatgggggaaacccctgaagcgagcgcg 373  
 Db 301 ACGGGAGGACAGATTAAAGATTGCTCAATGAGGGGGAACCTGAAAGCAGCAGCGCGC 360  
 QY 374 gtgaacgaagaaggtctctcgtatgtlaaagttcatagca-ggaaaaaataagcaagat 432  
 Db 361 GTGAACGATGAAGGCTTTCGGATTGTAAAGTTAAAGCAGGGAATAAGCAGCGAT 420  
 QY 433 gtgatgagtgtaacctgcttaagcaacggcgtactaactgctcagcaacggcggttaacg 492  
 Db 421 GTGATGATGTACTGCTTCAAGCAGCGGCTAAGTACTAGTGCACACCGCGGTAAATCG 480  
 QY 493 tatgttgaaagcgttctgtcgaatacatgtggcgtlaaagggtgcgtlaagcggtatgtaa 552  
 Db 481 TATGGTCAAGCGTTGTTCCGAATCATTTGGCGCTAAAGGTGCGTAGCGGACATGTAAAG 540  
 QY 553 tcaagtgtagaaactcgtcggtcacaaccgtgtgcctcgtcaactgaatacaaatcgtga 612  
 Db 541 TCAGGTGTGAATACTCGCGGCTCAACTCGACCTGCACTTGAACATGTATGTGTGGAGT 600  
 QY 613 ttgggaagggcaagtgtgaattccaggtgttagcgtgtgaattcgtatgatatctgaggaac 672  
 Db 601 TTGGGAGAGGCAAGTGAATTCACAGTGTAGCGGGAATCGTAGATATCTGAGAGAAC 660  
 QY 673 accagttgcgaagcgacttctgtcgtcaaaaactgacgctgaggaacgagcggtggtla 732  
 Db 661 ACCAATGCGGAAGGCGACTTGTGGCTTAAACTATGACCTGAGGACAGAAAGCGGGTA 720  
 QY 733 gtaaacggatagatgatacccggtatacccaagcctaaagttgtctaccagttgttgg 792  
 Db 721 GTGAACGGGATTAAGTATACCCCGTAAATCCAGCCCTTAAACGTGTCTCAAGTGTGGG 780  
 QY 793 ggtttaaaccctcaataacgaacccaacggatgaagttagcgagcgtggggaactagtcg 852  
 Db 781 GGTTTTAACCTCAGTAACGAACCTTAACGATTAAGTAGACCGCTGGGAGCTATAGCTCG 840  
 QY 853 caagagtgaaactcaagaagatlgacggtgtcgcacaaagcgtgtgagcatgtgttta 912

```

|||||
Db 841 CAAGAGTGAACCTCAAGGAATTGACGGGGTCCGCACAAAGCGGTGAGACATGCTTGA 900
Oy 913 atctgatgtaccaccaaaactcaactccttggttgacatgatacttgatcaatgtaga 972
Db 901 ATTGCGTGTAGCGCGCAAAAACCTCAGCTAGGCTTGACATGAGTGAATCATGTAGAGAT 960
Oy 973 atctgagccttcgggagagatcacaagtgctgaagttctgcagacgtctgtctgag 1032
Db 961 ACATGAGCCTTGGGGCGCGCTTACAGAGTGTGTCATGTTGTGTCAGCTGCTGTGAG 1020
Oy 1033 atcttggtttaagctcccgcaagcgcaaccctcctcgtatgttcta-ccttaagttg 1091
Db 1021 ATGTTGGGTTAAGTCCCGCAAGAGCGCAACCTCATTATGTTGCCATTCATTCAGTTG 1080
Oy 1092 ggcactggtgacgaactcgcgtgtgacaacccggaggaagcggggagtgacgtcaact 1151
Db 1081 GGCACCTCGTAAGAACCTGCCGGTGACAAACCGAGAGAGCGGGGATGACGTCAAATCCCT 1140
Oy 1152 catgagccttatgtccaggcgcaacacagtgctacaaatgagcagatacagaggtgcgcaa 1211
Db 1141 CATGGCCTTATGTCTAGGGCAACACACGTGCTACATAGGCGGTACAAAGGTTAGCCAA 1200
Oy 1212 ctcgcaagcgagagctaatctcttaaaagtcggtccagttcgatgggtctgcaactc 1271
Db 1201 CTCGGCAGGGGAGGTAAATCTCAAAAAGCCGGTCCAGTTGGATGGAGTGTGCAATCTC 1260
Oy 1272 gaccocatagaatcgcgaatcgcgtgaagtaacgcggaatcgcgagtgatgaatcgttc 1331
Db 1261 GACTCCATTAAGTCGGAATCGCTAGTAATCGCGGATCAGCATCTCCGGTGAATACGTTTC 1320
Oy 1332 ccggaactctgtacacacgcgcctcacaccacctgagtgaggagaccccgaaatggtctc 1391
Db 1321 CCGGACTGTGTACACACGCCCGCTCACCACCTGAGTGGGAGACACCCGAATGCTCTT 1380
Oy 1392 tcttaacgtaagagacagactactaagtgaaactcgtaag 1436
Db 1381 TCACAACCCGCAAGAGAGACAGACTAAGGTGAATCGTGAAGG 1425

RESULT 11
WB4616S 1486 bp DNA BCT 18-FEB-1995
LOCUS Leptospira weilii Sarmin 16S rRNA gene, partial sequence.
DEFINITION U12673
ACCESSION U12673
VERSION 012673.1 GI:558932
KEYWORDS
SOURCE
ORGANISM
Leptospira weilii.
Bacteria; Spirochaetales; Leptospiraceae; Leptospira.
REFERENCE
1 (bases 1 to 1486)
AUTHORS Ralph D. and McJelland, M.
TITLE Phylogenetic evidence for horizontal transfer of an intervening
sequence between species in a spirochete genus
JOURNAL J. Bacteriol. 176 (19), 5982-5987 (1994)
MEDLINE 95014031
REFERENCE 2 (bases 1 to 1486)
AUTHORS McJelland, M.
TITLE Direct Submission
JOURNAL Submitted (25-JUL-1994) Michael McJelland, California Institute of
Biological Research, 11099 North Torrey Pines Road, La Jolla, CA
92037, USA

FEATURES
source
1..1486
/organism="Leptospira weilii"
/strain="serogroup Icterohaemorrhagiae, subserogroup
Sarmin, serovar sarmin, strain Sarmin"
/isolate="WB46"
/db_xref="taxon:28184"
<1..1486
/product="16S rRNA"
BASE COUNT 384 a 332 c 452 g 315 t 3 others
ORIGIN

```

```

Query Match      84.8%: Score 1251.8; DB 1; Length 1486;
Best Local Similarity 94.0%: Pred. NO. 5.9e-12;
Matches 1375; Conservative 0; Mismatches 80; Indels 7; Gaps 7;

Oy 1 gatcatggtcagaactcaacgcctggtcggtcgtcttaacaatgcaagtcgagcggtag 60
Db 14 GATCCCTGCTCAGAACTAACGCTGGCGGGCGCTTAAACATGCAAGTCAACGGGATG 73
Oy 61 caatactagcggcgaaacgggtgaglaaacglt-gglaactcctcgcagtcctggagla 119
Db 74 CAATATCCAGCGCGCAACCGGTGAGTAAACAGTGGGTAACTCTCCCAAGTCTGGAGTA 133
Oy 120 actttccgaaagaaagcctaataccgataagtcctgttgatcaacaagtttgatagta 179
Db 134 ACTTTCGAAAGGGAGCTTAATATCTGATATGCCAGAGTCATATGATTTTTCGGGTA 193
Oy 180 aagattatctgttgagatgagcccgccgacatagctatgttgatgagtaatggtctc 239
Db 194 AAGATTATTGCTCGSAGCTGAGCCCGCCGCGCATTTAGTGTGTGAGTATGCTC 253
Oy 240 accaagcgacgacgtgtagcggcctgagaggtgtccggtccacaa tgaactgagaca 299
Db 254 ACCAAGCGACGATCGTGTAGCCGCTGAGAGGCTGTTCGCCACAAATGGAATCTGAGACA 313
Oy 300 cgttcatactcctcagggagggcagagttlaagaatctgtcgaatgggggaaacctga 359
Db 314 CGGTCCATACT-CTACGGAGGAGCAGATTAAGATCTTGCTCAATGGGGGAAACCTTA 372
Oy 360 agcagcgacgcgcgtgaaacgaagaagtcctcgatctgaaagttcatlaagca-ggaa 418
Db 373 AGCAGGAGCGCGCGGTGAAGATGAAGATGCTTCGATTTAAAGTTCAATAGCAGGAA 432
Oy 419 aaataagcgacaaatgtatgatagttaactgtcctaaagcaggcgttaactggtgcagca 478
Db 433 AATAAGCAGCAATGTGATGATGTGATCTGCTCAATTAATCAACCGGTAACTACGTCCAGCA 492
Oy 479 ggcgcgttaactcgtatgtgtgcaagcgtgtgttcggaatacttgggggttaaagggtgcgta 538
Db 493 GCCGCGTAAATAGTATGTGTGCAAGCGCTGTTCGGAAATCATTTGGCGGTAAAGGTCGTA 552
Oy 539 ggcggaatctgtaagtcaggtgtgtaaaactcgcggtcctcaacccgtggtcgtcaacttgaac 598
Db 553 GCGGAGCATGTAAAGTCAGGTGTC-AAAATGCGGGGTCAACTCGCACCTTGCACTTGAAC 611
Oy 599 tacaagtcctggagcttctggagaggaagtcgaatgtccaaagtgtgagtcgtgagtcgtg 658
Db 612 TATGTGTCTGGAATTTGGAGAGGCAAGTGGAATTCACAGGTGTAGCGGTAATGCGTGA 671
Oy 659 atactggaaggaacacacagctggcggaagcgactgtgtgtcctcaaaactgacgtgagca 718
Db 672 ATATCTGGAGGAACACCAAGTGGCGAAGGCGACTGTGCTTAAACTGACCTGAGGCA 731
Oy 719 cgaagcgtgtggtagtaaaacggagatagatacccggtatccacgcgtcctaaacgtgtgc 778
Db 732 CGAAACGCTGGGTAGTGAACGGGATAGATATACCCCGTATCCACGCCCTAAACGTTGTC 791
Oy 779 taccagttgtgggttttaaccctcagtaagcaacttaaggaatgaagttgtagacgctc 838
Db 792 TACCACTGTGTGGGGTTTTAACCTCAGTACGAA-CTAACGGATTAAGTGAACCGCT 850
Oy 839 gggagactatgctcgcaagag tgaactcaaaaggaatctgacgggggtcgcgacaaagcggtc 898
Db 851 GGGGACTATGCTCGCAAGAGTGAACCTCAAGGAATTTGACGGGGGTNNNACACAGCGGTG 910
Oy 899 gagcagtggtttaacttcagatgatacccaaaaacctcaactcgtgggtctgacatgatactc 958
Db 911 GACCATGTGTTAATTCATGATGATGACGCAAAAACCTCACTAGCTTACATGAGAGTGG 970
Oy 959 aatcagtgagagataatagacgttcggcgagatcaacagggtgtctgacagtggtgtcgtca 1018
Db 971 AATTATGTAGATTAACATGAG-CTTCGGGCGCGCTTACAGAGTGTCTCATGTTGCTGCA 1029

```



[illegible]

RESULT	13
LOCUS	LSU12672
DEFINITION	Leptospira santarosai L781 16S rRNA gene, partial sequence.
ACCESSION	U12672
VERSION	U12672.1 GI:558931

SOURCE  
ORGANISM  
Leptospira santarosai  
Leptospira santarosai  
Bacteria; Spirochaetales; Leptospiraceae; Leptospira.

REFERENCE	AUTHORS	TITLE	JOURNAL
1 (Pages 1 to 1494)	Ralph, D. and McCelland, M.	Phylogenetic evidence for horizontal transfer of an intervening sequence between species in a spirochete genus	J. Bacteriol. 176 (71), 5982-5987 (1994)

**MEDLINE** 95014031  
**REFERENCE** 2 (bases 1 to 1494)  
**AUTHORS** McClelland, M.  
**TITLE** Direct Submision  
**JOURNAL** Submitted (23-JUL-1994) Michael McClelland, California Institute of  
Biological Research, 11099 North Torrey Pines Road, La Jolla, CA  
92037, USA

FEATURES	Location/Qualifiers
source	1. .1494 ;

```

/organism="Leptospira santarosai"
/strain="serogroup Tarassovi, subserogroup Atlantae
serovar atlantae, strain L781"
/isolate="SA39"
/db_xref="taxon:20183"
<!.1494
rRNA

```

	/product="16S rRNA"				
BASE COUNT	383 a	332 c	453 g	311 t	15 others
ORIGIN					

Query Match	83.9%	Score 1239.8;	DB 1;	Length 1494;
Best Local Similarity	93.2%;	Pred. No. 8.8e-12;		
Matches 1363; Conservative	0;	Mismatches 92;	Indels 7;	Gaps 7;

[illegible]

Db	202	AAGTTTCATTGCTCGGAGGATGACCCCGGCTCCGATTAGCTACTGTTGGTATGAGTAAAGGCTC	261
Oy	240	accaagcgacgacatcvgttagccgctgtagaggtgtgccgccaatlgnaactlgagaca	299
Db	262	ACCAAGCGACGATCGGTATACCGGCTGAGAGGGTGTGGCGCACAAATGGAATGAGACA	321
Oy	300	cggtccacacccctacccgagggagcagaagttaaagaactctgtcccaatlgggggaaaccctgca	359
Db	322	CGGTCCACTACT - CTACGGGAGGCAAGCATTTAAATCTTGCTCAATGGGGGAAACCTCTGA	380
Oy	360	agcagcgacgcgcggtlgaaagaaaggttccttgcgatctgtaaagtlctatagca -ggaa	418
Db	381	AGCAGCGACGCGCGGTGAACGATGAAAGTCTTGATTCGATTGAAGTTCAATAAGCAGGAA	440
Oy	419	aaataagcagcgaatgtgatagtgtatccctgtccctaaagcccggtcaactaactggtccagca	478
Db	441	AAATTAAGCAGGATGTATATATGTGATCTGCTAAAGCAACCGGCTPACTACTGTCCACACA	500
Oy	479	gcccgcgttaaacgtatgtgtgcaagcgtgtctgcgatacattgggcgtlaaaggtgtcga	538
Db	501	GCCCGCGTAATACCTATGTTGTCGAACAGCGTCTTGCGAATCATTTGGCGTAAAGGTGCGTA	560
Oy	539	ggcggattgtgaagtcaggtgtgtaaaactgtcggtgtcgaacccgtggtcgtcaacttgaac	598
Db	561	GGCGGACATGTAAATCTCAGGCTG - AAACCTGCGGGCTCAACTCGCAGCCTGCACCTTGAAC	619
Oy	599	tacaagctcgtgaagtttggagaggaagctggaattccgaagtgttaacggtgaaatcgtgaa	658
Db	620	TATGTTCTGGAGTTTGGGAGGCAAGTGGAATTCAGAGTGAAGGAGTGAATCGATG	679
Oy	659	atattctggaggaacaacagctgtgcgaagcgactctgtgtctcaaaactlgacgtlgagga	718
Db	680	ATATCTGGAGGAACACCAAGTGGGGAAGGCACTTGCTGGCTTAAACTGACGCTGAGCA	739
Oy	719	cgaagcggtgtgtgtatgataacgggattagatacccggttaatccaaogccctaaagttgc	778
Db	740	CGAAAGCGTGGGTAGTGAACGGGATTAGATCCCGGTATCCAGCCCTTAACCTTTCTC	799
Oy	779	taccagttgttgggggtcttlaaccctgaagaaagaaactcaacgaattaaagtagaacgcct	838
Db	800	TACCGATTGTTGGGGGTTTAACCCCTAGTAACGA - CTAAAGGATTAGTAGACCGCT	858
Oy	839	ggggactatgctgcgcaagagtgaactcaaaaggaattgacgggggttcgcgacaagcgtgt	898
Db	859	GGGGCTATGCTCCCAAGAGTGAACCAAGGAATTGACGGGGGTTNNCACAAACGGGTG	918
Oy	899	gagcagtgtgtttaaattcgtatgaaccccaaaactccacgtgggtcttgaactgtgactgt	958
Db	919	GAGCATGTGGTTTAATTTCGATGATACCGCAAAAACCTCACCTTAGGTTGACTGGAAGTGG	978
Oy	959	aatcatgtagagaaatalgagcccttcggcgagaaattcaacaggtgtctgacatgtgttcgtca	1018
Db	979	AATCATGTAGAGATACATGAG - CTTCGGGGCGGCTTCACAGGTGCTCAGTGGTTGTCFCA	1037
Oy	1019	gtctgtgtctgtgagatgttgggtttaaagtcctgcgcaacgagcgcaacccctatcgtatgtgt	1078
Db	1038	GCTCGTGTGCGAATGTTGGGTTAAATCCCGCAACGACCGCAACCTCACCTTATGTGG	1097
Oy	1079	cta -ccctaaagtctggagcagtgaacatgtgcggttgacaacacccgaggaagcgggga	1137
Db	1098	CCATCATTTAATTGTGGCACTCGTAGGAACCTGCCGGTGACAAACCGAGAAAGCGGGGA	1157
Oy	1138	tgacgtcaaatccctcatgaccttatgttccaggggcacaacacgtgtactaatgtgcgata	1197
Db	1158	TGACGTTNNNNNNNNNNNGCCTTATGTCTATGCGCAACACAGCTGCTACAAATGGCGGTA	1217
Oy	1198	cagaggtgtcgccaactctgcaagagggaggtcaatctctaaaggtcgggtcccaattctggaatt	1257
Db	1218	CAAAGGTAGCCAACCTGCGCAGGAGGAGCTAATCTCAAAAAGCCGGTCCCAATTTGGAAAT	1277
Oy	1258	gggggtctgcaactctgaccccatggaagtcggaatcgcgtatgaatcgcggaatcagcatgcgt	1317
Db	1278	GGAGTCTCACTGACTTCATGTAAGTGGAAATCGCTACTAATCGCGATCATAGCAATGGCCG	1337





```
RESULT 15
L16SRN      1396 bp      mRNA      BCT      24-AUG-1993
LOCUS       L16SRN
DEFINITION  L16SRN
ACCESSION   Z12817
VERSION     Z12817.1 GI:44003
KEYWORDS    16S ribosomal RNA; ribosomal RNA.
SOURCE      Leptospira interrogans.
ORGANISM    Bacteria; Spirochaetales; Leptospiraceae; Leptospira.
REFERENCE   1 (bases 1 to 1396)
AUTHORS     Bryden, J., Gatehouse, L., Gatehouse, J. A. and Hooke, J. V.
TITLE       Automated PCR cycle sequencing of 16S ribosomal RNA genes
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 1396)
AUTHORS     Hooke, J. V.
TITLE       Direct Submission
JOURNAL     Submitted (19-JUN-1992) Hooke, J. V., Public Health Laboratory
            Service, Leptospira Reference Laboratory, Stonebow Road, Hereford,
            Herefordshire, United Kingdom, HR1 2ER
            Location/Qualifiers
FEATURES             location=1396
source               /organism="Leptospira interrogans"
                    /strain="lepterohaemorrhagiae RGA"
                    /db_xref="taxon:173"
                    1..1396
                    /standard_name="16S ribosomal RNA"
BASE COUNT      365 a      319 c      418 g      294 t
ORIGIN
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209
2210
2211
2212
2213
2214
2215
2216
2217
2218
2219
2220
2221
2222
2223
2224
2225
2226
2227
2228
2229
2230
2231
2232
2233
2234
2235
2236
2237
2238
2239
2240
2241
2242
2243
2244
2245
2246
2247
2248
2249
2250
2251
2252
2253
2254
2255
2256
2257
2258
2259
2260
2261
2262
2263
2264
2265
2266
2267
2268
2269
2270
2271
2272
2273
2274
2275
2276
2277
2278
2279
2280
2281
2282
2283
2284
2285
2286
2287
2288
2289
2290
2291
2292
2293
2294
2295
2296
2297
2298
2299
2300
2301
2302
2303
2304
2305
2306
2307
2308
2309
2310
2311
2312
2313
2314
2315
2316
2317
2318
2319
2320
2321
2322
2323
2324
2325
2326
2327
2328
2329
2330
2331
2332
2333
2334
2335
2336
2337
2338
2339
2340
2341
2342
2343
2344
2345
2346
2347
2348
2349
2350
2351
2352
2353
2354
2355
2356
2357
2358
2359
2360
2361
2362
2363
2364
2365
2366
2367
2368
2369
2370
2371
2372
2373
2374
2375
2376
2377
2378
2379
2380
2381
2382
2383
2384
2385
2386
2387
2388
2389
2390
2391
2392
2393
2394
2395
2396
2397
2398
2399
2400
2401
2402
2403
2404
2405
2406
2407
2408
2409
2410
2411
2412
2413
2414
2415
2416
2417
2418
2419
2420
2421
2422
2423
2424
2425
2426
2427
2428
2429
2430
2431
2432
2433
2434
2435
2436
2437
2438
2439
2440
2441
2442
2443
2444
2445
2446
2447
2448
2449
2450
2451
2452
2453
2454
2455
2456
2457
2458
2459
2460
2461
2462
2463
2464
2465
2466
2467
2468
2469
2470
2471
2472
2473
2474
2475
2476
2477
2478
2479
2480
2481
2482
2483
2484
2485
2486
2487
2488
2489
2490
2491
2492
2493
2494
2495
2496
2497
2498
2499
2500
2501
2502
2503
2504
2505
2506
2507
2508
2509
2510
2511
2512
2513
2514
2515
2516
2517
2518
2519
2520
2521
2522
2523
2524
2525
2526
2527
2528
2529
2530
2531
2532
2533
2534
2535
2536
2537
2538
2539
2540
2541
2542
2543
2544
2545
2546
2547
2548
2549
2550
2551
2552
2553
2554
2555
2556
2557
2558
2559
2560
2561
2562
2563
2564
2565
2566
2567
2568
2569
2570
2571
2572
2573
2574
2575
2576
2577
2578
2579
2580
2581
2582
2583
2584
2585
258
```

---



GenCore Version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 01:18:20 : Search time 401.91 Seconds  
(without alignments)  
3150.629 Million cell updates/sec

Title: US-09-380-826A-1

Perfect score: 1477

Sequence: 1 gataatgacagactaac.....ccgtaaatgcattctcag 1477

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: N.Geneseq-1101.\*  
2: /SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq/NA1986.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT.\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1477	100.0	1477	19	AAV58896
2	827	56.0	1532	22	AAV58896
3	824.8	55.8	1528	22	AAV58896
4	815.6	55.2	1529	22	AAV58896
5	815	55.2	1533	22	AAV58896
6	814.6	55.2	1535	20	AAV58896
7	814.4	55.1	1485	22	AAV58896
8	813.4	55.1	1512	20	AAV58896
9	811.6	54.9	1529	20	AAV58896
10	810.2	54.9	1508	20	AAV58896
11	810.2	54.9	1508	21	AAV58896

12	810	54.8	1529	20	AAV58896
13	808.4	54.7	1529	20	AAV58896
14	808.4	54.7	1529	20	AAV58896
15	807.8	54.7	1513	20	AAV58896
16	805.4	54.5	1521	16	AAV58896
17	805.4	54.5	1526	16	AAV58896
18	805.2	54.5	1460	22	AAV58896
19	805	54.5	1528	20	AAV58896
20	805	54.5	1535	20	AAV58896
21	803.8	54.4	1536	17	AAV58896
22	802.8	54.4	1477	20	AAV58896
23	802.4	54.3	1627	22	AAV58896
24	801.8	54.3	1512	20	AAV58896
25	801.6	54.3	1555	17	AAV58896
26	801.6	54.3	1555	19	AAV58896
27	801.4	54.3	1450	21	AAV58896
28	801.2	54.2	1526	22	AAV58896
29	800.8	54.2	3169	22	AAV58896
30	800.8	54.2	3308	22	AAV58896
31	800.8	54.2	3657	22	AAV58896
32	800.4	54.2	1532	13	AAV58896
33	798	54.0	1460	22	AAV58896
34	797.8	54.0	1460	22	AAV58896
35	797.6	54.0	3821	22	AAV58896
36	796.4	53.9	1506	20	AAV58896
37	795.6	53.9	1532	15	AAV58896
38	794.8	53.8	1446	17	AAV58896
39	794.8	53.8	1535	19	AAV58896
40	793.8	53.7	1459	22	AAV58896
41	793.6	53.7	1535	20	AAV58896
42	792	53.6	349980	22	AAV58896
43	792	53.6	349980	22	AAV58896
44	792	53.6	349980	22	AAV58896
45	792	53.6	349980	22	AAV58896

## ALIGNMENTS

RESULT 1	
ID	AAV58896
AAV58896	standard; DNA; 1477 BP.
AC	AAV58896;
XX	
DT	20-JAN-1999 (first entry)
XX	
DE	L. fainei nucleotide sequence.
XX	
KW	Infection; pathogenic Leptospira; protective immunity; therapy;
KW	diagnosis; ss.
XX	
OS	Leptospira fainei.
XX	
PN	W09840099-A1.
XX	
PD	17-SEP-1998.
XX	
PF	06-MAR-1998; 98WO-AU00145.
XX	
PR	07-MAR-1997; 97AU-0005494.
XX	
PA	(AGRI-) AGRIC VICTORIA SERVICES Pty LTD.
PA	(PIGR-) PIG RES & DEV CORP.
XX	
PI	Chappel RJ;
XX	
DR	WPI; 1998-520791/44.
XX	
PT	New isolated pathogenic Leptospira bacterium - useful for, e.g
PT	developing products for conferring protective immunity, and for
XX	prophylactic or therapeutic treatment
XX	

PS Claim 15; page 69-70; 94pp; English.

This sequence represents a *Leptospira* DNA sequence isolated from the pathogenic *Leptospira* (LS) bacterium of the invention. The bacterium belongs to serogroup *Hurstbridge* or *serovar hurstbridge* or the species *L. falnei*. The LS bacterium can be used for conferring protective immunity against pathogenic LS bacteria in humans or animals. The bacteria can also be used for prophylactic or therapeutic treatment and infections. The DNAs and antibodies may also be used for detection and diagnosis of past or present LS infection.

Sequence 1477 BP; 390 A; 334 C; 439 G; 314 T; 0 other;

Query Match	100.0%	Score 1477	DB 19	Length 1477
Best Local Similarity	100.0%	Pred. No. 0		
Matches 1477	0	Mismatches 0	Indels 0	Gaps 0

Oy	1	gataatggcgaagaactaaagcctgvggagcgctcttaacaatgcaagctcgvgggtag	60
Db	1	gataatgctcagaactaaagcctgvggagcgctcttaacaatgcaagctcgvgggtag	60
Oy	61	caatacctagcvgvgaaacvgvgtagtaacaacgtggtaaactctctccagatctvggataa	120
Db	61	caatacctagcvgvgaaacvgvgtagtaacaacgtggtaaactctctccagatctvggataa	120
Oy	121	cttcgcgaagggaaagctaaataacgggaatgctcgtgtggaatcaaaaagattgataagtaa	180
Db	121	cttcgcgaagggaaagctaaataacgggaatgctcgtgtggaatcaaaaagattgataagtaa	180
Oy	181	agatttatgtcttgagaatatagaccgcgcgcgcgaattagctagcttggtagaattgctca	240
Db	181	agatttatgtcttgagaatatagaccgcgcgcgcgaattagctagcttggtagaattgctca	240
Oy	241	ccaaagcgaacgaatcggtgaacccgctcgagagaggtgtgtccggccacaatggaactgagaaac	300
Db	241	ccaaagcgaacgaatcggtgaacccgctcgagagaggtgtgtccggccacaatggaactgagaaac	300
Oy	301	ggtccatatctctccgcggagaggaacgaatgaagaattctgtcaatgvgvgaaacctgaa	360
Db	301	ggtccatatctctccgcggagaggaacgaatgaagaattctgtctcaatgvgvgaaacctgaa	360
Oy	361	ggaagcgaacgcgcgtgagacggaagaggtctctcgagatgtgaagaattcatatgacaggaata	420
Db	361	ggaagcgaacgcgcgtgagacggaagaggtctctcgagatgtgaagaattcatatgacaggaata	420
Oy	421	ataagcagcaaatgtgatagtatgatactgcgtcaataagcaacggctcaactaagtgccagcagc	480
Db	421	ataagcagcaaatgtgatagtatgatactgcgtcaataagcaacggctcaactaagtgccagcagc	480
Oy	481	ggcggtaatacgtatgagtgacgaacgctgttctcggaatacatatgvggcgtcaaaaggttggtatgg	540
Db	481	ggcggtaatacgtatgagtgacgaacgctgttctcggaatacatatgvggcgtcaaaaggttggtatgg	540
Oy	541	cggatttgtaaatcgaagctgtgaaaaactvgcgggtccaacccgtgctgcacttgaaacta	600
Db	541	cggatttgtaaatcgaagctgtgaaaaactvgcgggtccaacccgtgctgcacttgaaacta	600
Oy	601	caagctctgagatcttgvggagaggaacgaatggaatctccaaagtgtgacggtgaaatgctgatact	660
Db	601	caagctctgagatcttgvggagaggaacgaatggaatctccaaagtgtgacggtgaaatgctgatact	660
Oy	661	atctgagagaaacaacgaatgvcgaagvcgaactctgtgctctcaaaaactgacgcgtgagacag	720
Db	661	atctgagagaaacaacgaatgvcgaagvcgaactctgtgctctcaaaaactgacgcgtgagacag	720
Oy	721	aaagcgttggtgtagtaaaacgggattagatacccggtcaatccaaagccctaaagctgttcta	780
Db	721	aaagcgttggtgtagtaaaacgggattagatacccggtcaatccaaagccctaaagctgttcta	780
Oy	781	ccaagctgtctgggggttttaaacctcagtaagaacactcaagaagattgaagtagagccgctgg	840
Db	781	ccaagctgtctgggggttttaaacctcagtaagaacactcaagaagattgaagtagagccgctgg	840

QY	841	ggacatctgctcgcaagagtggaacatccaaagatgtgaagggggggtccgcaaaacggttga	900
Db	841	ggactaagctcgcaagagtggaacatccaaagatgtgaaggggggtccgcaaaacggttga	900
QY	901	gcatgtgttaattctgatgatataccccaacacactacccttggtcttgacatgtaactgaa	960
Db	901	gcatgtgtttaaattctgatgatataccccaacacactacccttggtcttgacatgtaactgaa	960
QY	961	tcatgtagagatatatagagctctcgggcgagattcacagggtgctcatggtgtgtctaacg	1020
Db	961	tcatgtagagatatataagagctctcgggcgagattcacagggtgtgtcatggtgtgtctaacg	1020
QY	1021	tcggtgttgagatgtgttgggttaagatcccgcaacagagcgcaaccctatcgtatgtgtct	1080
Db	1021	tcggtgtgtgagatgtgttgggttaagatcccgcaacagagcgcaaccctatcgtatgtgtct	1080
QY	1081	accttaagttgggcaacttgtaagaactctcggttgacaacacccggagggagagcggtgtga	1140
Db	1081	accttaagttgggcaacttgtaagaactctcggttgacaacacccggagggagagcggtgtga	1140
QY	1141	cgtaaatccctcagaggtcttatgtctcagaaggccacacacgttgctacaaatggccgatacag	1200
Db	1141	cgtaaatccctcagaggtcttatgtctcagaaggccacacacgttgctacaaatggccgatacag	1200
QY	1201	agggtcgccaactctcgcaagagggagctaatctcttaaaatctgtgtccaggtccgatttggg	1260
Db	1201	agggtcgccaactctcgcaagagggagctaatctcttaaaatctgtgtccaggtccgatttggg	1260
QY	1261	gtctgcgaactcgaccatctgaatctgaatctgtaatactcggtatcagatcagatcgccgg	1320
Db	1261	gtctgcgaactctcgaccatctgaatctgaatctgtaatactcggtatcagatcagatcgccgg	1320
QY	1321	tgaataagttcccgagacctgtgtacacacgcgcgtacacacacccctggaatggggaagacc	1380
Db	1321	tgaataagttcccgagacctgtgtgtacacacgcgcgtacacacacccctggaatggggaagacc	1380
QY	1381	gaagtggtgctctgtttaacccgtaagagacagactactaagtgtaaaatctgttaaaagggtct	1440
Db	1381	gaagtggtgctctgtttaacccgtaagagacagactactaagtgtaaaatctgttaaaagggtct	1440
QY	1441	gaagtcgtaacaaggtaccgttaaatcgatctctcgag	1477
Db	1441	gaagtcgtaacaaggtaccgttaaatcgatctctcgag	1477
RESULT 2			
AAAF76234			
ID	AAAF76234 standard; DNA; 1532 BP.		
XX			
AC	AAAF76234;		
XX			
DT	05-JUN-2001 (first entry)		
XX			
DE	Cycloclasticus pugetii 16S rRNA gene sequence, SEQ ID NO:1.		
KX	16S rRNA gene; ribosomal RNA gene; petroleum degrading bacterium;		
KW	hydrocarbon; detection; quantitation; identification; oil spill;		
KM	polluted environment; environmental decontamination; ds.		
OS	Cycloclasticus pugetii.		
XX			
PN	MO200114587-A1.		
XX			
PD	01-MAR-2001.		
XX			
PF	24-AUG-2000; 2000MO-JP05711.		
XX			
PR	25-AUG-1999; 99UP-0237818.		
XX			
PA	(AGEN ) AGENCY OF IND SCI & TECHNOLOGY.		
PA	(NISH-) NISHIMATSU CONSTRUCTION CO LTD.		
PA	(NYKL-) NYK LOGISTICS TECHNOLOGY INST CO LTD.		

RESULT	2	
AAAF76234		
ID	AAAF76234	standard; DNA; 1532 BP.
XX		
AC	AAAF76234;	
XX		
DT	05-JUN-2001	(first entry)
XX		
DE	Cycloclasticus pugetii 16S rRNA gene sequence, SEQ ID NO:1.	
XX		
KM	16S rRNA gene; ribosomal RNA gene; petroleum degrading bacterium	
KM	hydrocarbon; detection; quantitation; identification; oil spill;	
KM	polluted environment; environmental decontamination; ds.	
XX		
OS	Cycloclasticus pugetii.	
XX		
PN	WO200114587-A1.	
XX		
PD	01-MAR-2001.	
XX		
PF	24-AUG-2000; 2000WO-JP05711.	
XX		
PR	25-AUG-1999; 99JP-0237818.	
XX		
PA	(AGEN ) AGENCY OF IND SCI & TECHNOLOGY.	
PA	(NISH-) NISHIMATSU CONSTRUCTION CO LTD.	
PA	(NYKL-) NYK LOGISTICS TECHNOLOGY INST CO LTD.	

XX Maruyama A, Higashihara T, Ishiwata H, Fujita T;  
 XX WPI: 2001-218458/22.  
 DR  
 XX  
 PT Detection and quantitation of microorganism having specific function  
 PT and its gene from natural environment, with identification of  
 PT petroleum-digesting bacterium cycloclasticus for treating oil spillage  
 PT  
 XX  
 PS Claim 11: Page 40-41: 54pp: Japanese.

XX The invention relates to a method for detecting, identifying and  
 CC quantitating a microorganism which has a specific function, or DNA  
 CC therefrom, from the environment. The method comprises broadly classifying  
 CC the microorganism after serial dilution; amplifying and cloning specific  
 CC gene domains using extracted DNA as a template; examining the differences  
 CC between the gene domains cloned and determining their base sequences; and  
 CC positively identifying the microorganism from the base sequence data.  
 CC The invention also relates to a method to evaluate the function of such  
 CC an organism in various of environments, especially polluted  
 CC environments (particularly those contaminated with oil or harmful  
 CC chemicals) using migratory analysis. The methods enable the detection  
 CC and quantitation of a microorganism obtained from a polluted site in  
 CC order to identify strains capable of degrading petroleum and harmful  
 CC chemicals. Such organisms are particularly used for environmental  
 CC decontamination. The invention is particularly concerned with the  
 CC detection and quantitation of petroleum-degrading bacteria of the genus  
 CC Cycloclasticus. Cycloclasticus puegetii 16S rRNA gene sequences  
 CC (AA76334-AA76337) are specifically claimed, as are probes (AA76338-  
 CC AA76340) which hybridise to these sequences. The present sequence  
 CC is one of the claimed Cycloclasticus puegetii 16S rRNA (ribosomal  
 CC RNA) gene sequences.

Sequence 1532 BP: 392 A: 352 C: 473 G: 314 T: 1 other:

Query Match 56.0%; Score 827; DB 22; Length 1532;  
 Best Local Similarity 76.0%; Pred. No. 5,1e-247;  
 Matches 1138: Conservative 1; Mismatches 316; Indels 42; Gaps 8;

QY 1 gatactgctcgaactaagctgctgctgctcctaacaatgacgaagtcgagcggtag 60  
 DB 8 gatctgctcgaactaagctgctgctgctcctaacaatgacgaagtcgagcggtag 67  
 QY 61 caatc-----ctagcgcggaacgggtggttaacacgttgtaattctt 102  
 DB 68 gaatcgacgtctgctgacgagcggtcgaatgagcgagcggtggttaacacgttgtaattctt 127  
 QY 103 cctccgagctggtgaatacttcgaaagaaagaaactaatacggatagctctgtgatac 162  
 DB 128 ccgagtagtgggggaaacactcctctgaacgctgctaatacggatagctctgtgatac 185  
 QY 163 acaagattgatagttaagaatattctctggaatgagccgagcgccgattagtagt 222  
 DB 186 gaaagaacggggaacgttcgagccttgcttaatgtagctgacacaggggattagtagt 245  
 QY 223 tggtagtgtaagctcaccacaaagcgagatcgtagcgccctggaaggtgttcggcgcc 282  
 DB 246 tggtagtgtaagctcaccacaaagcgagatcgtagcgccctggaaggtgttcggcgcc 305  
 QY 283 acaatggaactagacagctcactactcctacgaggaagcagacagttaagaatctgtctc 342  
 DB 306 acacttggaactagacagctcactactcctacgaggaagcagacagttaagaatctgtctc 365  
 QY 343 aatgggggaaacccctgaacagcgagcgcggtgaacgagaaggtcctcgatgttaaa 402  
 DB 366 aatgggggaaacccctgaacagcgagcgcggtgaacgagaaggtcctcgatgtgttaaa 425  
 QY 403 gtccatagcgagaaataagcagcaatgtgatagttactctgcta----- 452  
 DB 426 gcaacttcagtaggggaaagtttaaggtaataaccccttagcgcccgagcttaacta 485

QY 453 -----aagcaccggttaactacgttcgacagcagccggttaactagctgtaacagcgt 506  
 DB 486 cagaagaagcaccggttaactacgttcgacagcagccggttaactagctgtaacagcgt 545  
 QY 507 tgttcggaactattgggctaaagggtcgtgagcggaatttgaagtcaggtgtgaaac 566  
 DB 546 taatcggaactattgggctaaagggtcgtgagcggaatttgaagtcaggtgtgaaac 605  
 QY 567 tggggtcgaacccgttggtctgcaacttgaaactaagaatcgtgaggttggggaggaag 626  
 DB 606 cccgggtcgaacccgttggtctgcaacttgaaactgagtagtggtgaggaaggaag 665  
 QY 627 tgaattccaggttgtagcggttaaatgctagatatctgaaagaacacagctggtgaag 686  
 DB 666 tgaattccaggttgtagcggttgaaatgctagatatctgaaagaacacagctggtgaag 725  
 QY 687 cgaattgctgctcaaaactgacgctgaaagcagaagcgttggtagtaaacggttagt 746  
 DB 726 cgtctctcgacccaactgacgctgaaagcgttggtagtaaacggttagt 785  
 QY 747 ataccgggttaactcagccctaaacgttgctacacagttgttgggttttaacctta 806  
 DB 786 ataccgggttaactcagccctaaacgttgctacacagttgttgggttttaacctta 845  
 QY 807 gtaacgaaccttaacgattaaatagacgctggtggtgactatctcgcaagatggaactc 866  
 DB 846 ggtggtgca-staacgaataaagttagacgctggtggtgactatctcgcaagatggaactc 904  
 QY 867 aaagaattgagcggtggttcgcaacagcgttgaagcattgtgatttaattcgatgatacc 926  
 DB 905 aaatgaattgagcggtggttcgcaacagcgttgaagcattgtgatttaattcgatgatacc 964  
 QY 927 caaaacctcaccctgggttgatcagatcga-atcagtagtagatatagaccttg 985  
 DB 965 gaagaaccttaaccttacccttgatcagatcga-atcagtagtagatatagaccttg 1024  
 QY 986 ggcagatt--cacaagtgctgcatggtgtgctgacagctcgtgctgtaagatgttggtta 1043  
 DB 1025 ggaaccttgatacaggtgtgcatggtgtgctgacagctcgtgctgtaagatgttggtta 1084  
 QY 1044 agtccgcaacagcggaacccctatcgtatgttgtaacc-ttaagtgtggacatggtac 1102  
 DB 1085 agtccgcaacagcggaacccctatcgtatgttgtaacc-ttaagtgtggacatggtac 1144  
 QY 1103 gaaactcgtggtgaacacacgggaggaagcggtggaatgacgttaaatccctatggcttta 1162  
 DB 1145 ggaactcgtggtgaacacacgggaggaagcggtggaatgacgttaaatccctatggcttta 1204  
 QY 1163 tgtccagggccacacacggtgctatacaatgacgaataggggttcgccaactcgcaagagg 1222  
 DB 1205 tggtagggctatacaacggtgctatacaatgacgaataggggttcgccaactcgcaagagg 1264  
 QY 1223 gaggtaactcttaaaagtcgttccacgttcggaattgggtgtcgcgaactcgaccctatga 1282  
 DB 1265 aagctaatcctcttaaaagtcgttccacgttcggaattgggtgtcgcgaactcgaccctatga 1324  
 QY 1283 gtcgcgaactcgtatgaatcgcgcatcagcatgcggtgtgaatacgttcgcaagctgtg 1342  
 DB 1325 gctcggaactcgtatgaatcgcgcatcagcatgcggtgtgaatacgttcgcaagctgtg 1384  
 QY 1343 acacacgcgcgttcacacacacgtgagtggtggaacacccgaagtgtgtcttgaacgcta 1402  
 DB 1385 acacacgcgcgttcacacacacgtgagtggtggaacacccgaagtgtgtcttgaacgcta 1443  
 QY 1403 aggaagcaactactaagttgaactcgttaaggggtggaatgtgtaaaagaagttacc 1459  
 DB 1444 ggaagcgctcaaccacttltgattcagtaactgtgagtggaagtcgtaacaaagttagc 1500

RESULT 3  
 AAF76235  
 ID AAF76235 standard; DNA: 1528 BP.  
 XX

AA76235;  
 05-JUN-2001 (first entry)  
 Cycloclasticus puegell 16S rRNA gene sequence, SEQ ID NO:2.  
 16S rRNA gene; ribosomal RNA gene; petroleum degrading bacterium;  
 hydrocarbon; detection; quantitation; identification; oil spill;  
 polluted environment; environmental decontamination; ds.  
 Cycloclasticus puegell.  
 W0200114587-A1.  
 01-MAR-2001.  
 24-AUG-2000: 2000WO-JP05711.  
 25-AUG-1999: 99JP-0237818.  
 (AGENCY OF IND SCI & TECHNOLOGY.  
 (NISH-) NISHIMATSU CONSTRUCTION CO LTD.  
 (NYKL-) NYK LOGISTICS TECHNOLOGY INST CO LTD.  
 Maruyama A, Higashihara T, Ishiwata H, Fujita T;  
 WPI: 2001-218458/22.  
 Detection and quantitation of microorganism having specific function  
 and its gene from natural environment, with identification of  
 petroleum-digesting bacterium cycloclasticus for treating oil spillage  
 Claim 11: Page 41-42: 54pp: Japanese.  
 The invention relates to a method for detecting, identifying and  
 quantitating a microorganism which has a specific function, or DNA  
 therefrom, from the environment. The method comprises broadly classifying  
 the microorganism after serial dilution; amplifying and cloning specific  
 gene domains using extracted DNA as a template; examining the differences  
 between the gene domains cloned and determining their base sequences; and  
 positively identifying the microorganism from the base sequence data.  
 The invention also relates to a method to evaluate the function of such  
 an organism in various of environments, especially polluted  
 environments (particularly those contaminated with oil or harmful  
 chemicals) using migratory analysis. The methods enable the detection  
 and quantitation of a microorganism obtained from a polluted site in  
 order to identify strains capable of degrading petroleum and harmful  
 chemicals. Such organisms are particularly used for environmental  
 decontamination. The invention is particularly concerned with the  
 detection and quantitation of petroleum-degrading bacteria of the genus  
 Cycloclasticus. Cycloclasticus puegell 16S rRNA gene sequences  
 (AA76234-AA76237) are specifically claimed, as are probes (AA76238-  
 AA76240) which hybridise to these sequences. The present sequence  
 is one of the claimed Cycloclasticus puegell 16S rRNA (ribosomal  
 RNA) gene sequences.  
 Sequence 1528 BP; 394 A; 343 C; 467 G; 323 T; 1 other:  
 Query Match 55.8%; Score 824.8; DB 22; Length 1528;  
 Best Local Similarity 76.0%; Pred. No. 2.4e-246;  
 Matches 1135; Conservative 1; Mismatches 318; Indels 40; Gaps 8;

128 taatagtgtggaacaacccgtgtgaaacacaggtataacgataatccctacgggcaaa 187  
 165 aagatttgataagtaagaattatgtcttgagagatgagcccgccgattagctgtg 224  
 188 agca--gggaccccttcggccttcgtcgtataatagatagcctatgtcgtatagctagt 245  
 225 gtaggttaagtgtctcccaagcgcagatcgtgtagccgctgagaggggtgtccggcac 284  
 246 gtaggttaagtgtctcccaagcgcagatcgtgtagccgctgagaggggtgtccggcac 305  
 285 aatggaacttgagacacaggtccatactctacggaagcagcttaagaatcttctctcaa 344  
 306 actcgagactgagacacaggtccatactctacggaagcagcttgggaatatgtgcacaa 365  
 345 tggggggaacccctgaagcagcgcgcgttgaaacgaagaagctcttcgattgaagt 404  
 366 tggaggaactctgtagcgaatgcgcgtgtgtgaaagaagcccttaagggtgttaagc 425  
 405 tcaatagcaggaanaataagcagcaatgtgatgtgtacctgtbcta----- 452  
 426 acttcagtagggaggaanaagtttaagttataacttaggcccttgagcttactacag 485  
 453 ---aagcaccgcttaactacgttccagcagccgcgttaatacgtatgtgcaagcgtgt 509  
 486 aagaagcaccgcttaactacgttccagcagccgcgttaatacgtatgtgcaagcgttaa 545  
 510 tcggaatcatttggcgttaaaagggtgtgctgtagcgatttgaatcaaggtgtgaaatcgc 569  
 546 tcggaatcatttggcgttaaaagggtgtgctgtagcgatttgaatcaaggtgtgaaatcgc 605  
 570 gggtcacaaccgtgtgcttgcacttgaaactaagaatctgtgagtttggagaagcagttg 629  
 606 gggtcacaaccgtgtgcttgcacttgaaactgattgaaactgtttagtagatgtgtg 665  
 630 aattccaggtgtgagcgttgaaatgtcgtatgtatcttgagaagaaacacagctggcgaagcga 689  
 666 aattcaggtgtgagcgttgaaatgtcgtatgtatcttgagaagaaacacagctggcgaagcga 725  
 690 ctgtgctcaaaaacagcgtctgagcagcagaagcgttgggttagtaaacgggattagata 749  
 726 ctctcttgaccacaacacagcgtcgtgagtgtaggaagcgttgggttagcaacggattagata 785  
 750 ccccggtatccacgcgccttaacagctgtgtctaccagttgttgggggttttaacctcagta 809  
 786 ccccggtatccacgcgccttaacagctgtgtgtctaccagttgttgggggttttaacctcagta 845  
 810 acgaaactaagcgttgaatgtagcgcgccttgggactatgtctcgcgaagaagtgaactcaaa 869  
 846 gtgca--staaacgcaataaagttgacccgcttgggggttagcgcgcgaagaagtgaactcaaa 904  
 870 ggaattgacgggggttcgcgacaagcgttggagcagtgtgttttaattcgtatatacccaa 929  
 905 tgaattgacgggggttcgcgacaagcgttggagcagtgtgttttaattcgtatatacccaa 964  
 930 aaactcaactgtgacatgtgatactga--ataatgtagagatatagaccttcgggc 988  
 965 gaaccttaacttaaccttgacataacagagaactttcttagaagataagttgtgtcctcgga 1024  
 989 agatt--cagaggtgtcgtacatgtgtgtcgtacagctcgtcgtgtgagatgttgggttaagt 1046  
 1025 actcgtatagcaggtgtcgtacatgtgtgtcgtacagctcgtcgtgtgagatgttgggttaagt 1084  
 1047 ccgcgaacgagcgaacccctatcgtatgtgtctac--ttaagtgtggcactgtgtacgaa 1105  
 1085 ccgcgaacgagcgaacccctatccttagtctgtacacatttagtttggcactcttaagag 1144  
 1106 actgcgggtgacaacacggaaggaagcgtgagatagcgtcaaatcctatctgtcctttagt 1165  
 1145 actgcgggtgacaacacggaaggaagcgtgagatagcgtcaaatcctatctgtcctttagt 1204  
 1166 ccagggcacaacagctgtctacaatgtgcgataacagaaggtgcgcaactcgcgaagaagggag 1225

Db	1205	gttaggctacacacgctgctacaaatggccggtacagagggccgcaactccgcgagatgaag	1204
Oy	1226	ctaatctctaaagttcggltccgattcggatttggtgtctgcgaactcgaaaccccatgaaatc	1205
Db	1265	ctaatcccttaaaagccggtctctagatccggattgcaatctgcgaactcgagctgatgaaact	1324
Oy	1286	ggaaatcgctgtaattcgccggaattagcatgcccgcggtggaataagttcccgagacttgtaca	1345
Db	1335	ggaaatcgctgtaattcgccggaattagcatgcccgcggtggaataagttcccgagacttgtaca	1384
Oy	1346	caccgcccgtacacacacccgagtggtggagagacccggaatggctctgtgttaacggtaaag	1405
Db	1385	caccgcccgtacacacccgagtggtgggtgttcgaaagaagatgg--gtaggctaaacttcggg	1442
Oy	1406	agacagacatctaaagtgaactcgtctaaaggggtgtgaagtcgttaacaaagtacc	1459
Db	1443	agcgccgtctacacactctgtgtatctacagactgtgggtgtgaagtcgttaacaaagtggc	1496

## RESULT 4

AAAF76236 standard; DNA; 1529 BP.

AA AF76236;

DT 05-JUN-2001 (first entry)

Cycloclasticus pugetii 16S rRNA gene sequence, SEQ ID NO:3.

16S rRNA gene; petroleum degrading bacterium; KW

polluted environment; environmental decontamination; ds.

05 Cycloclasticus pugetii.

PN WO200114587-A1.

01-MAR-2001. PD

PF 24-AUG-2000; 2000WO-JP05711.

PR 25-AUG-1999; 99JP-0237818.

PA (AGEN ) AGENCY OF IND SCI &amp;

PA (NYKL-) NYK LOGISTICS TECHNOLOGY INST CO LTD.

PI Maruyama A, Higashihara T, Ishiwata H, Fujii

DR WPI; 2001-218458/22.

PT Detection and quantitation of microorganism having

petroleum-digesting

XX XX

CC The invention relates to a method for detecting, identifying and  
CC quantitating a microorganism which has a specific function, or DNA  
CC therefrom, from the environment. The method comprises broadly classifying  
CC the microorganism after serial dilution; amplifying and cloning specific  
CC gene domains using extracted DNA as a template; examining the differences  
CC between the gene domains cloned and determining their base sequences; and  
CC positively identifying the microorganism from the base sequence data.  
CC The invention also relates to a method to evaluate the function of such  
CC an organism in various of environments, especially polluted  
CC environments (particularly those contaminated with oil or harmful  
CC chemicals) using migratory analysis. The methods enable the detection  
CC and quantitation of a microorganism obtained from a polluted site in  
CC order to identify strains capable of degrading petroleum and harmful  
CC chemicals. Such organisms are particularly used for environmental  
CC decontamination. The invention is particularly concerned with the

CC detection and quantitation of petroleum-degrading bacteria of the genus  
CC *Cycloclasticus*. *Cycloclasticus puetelli* 16S rRNA gene sequences  
CC *Cycloclasticus*. *Cycloclasticus puetelli* 16S rRNA gene sequences  
CC (AAAF6234-AAAF6237) are specifically claimed, as are probes (AAAF6238-  
CC AAAF6240) which hybridise to these sequences. The present sequence  
CC is one of the claimed *Cycloclasticus puetelli* 16S rRNA (ribosomal  
CC rRNA) gene sequences.

Sequence 1529 BP; 394 A; 343 C; 466 G; 320 T; 6 other;

Query Match	55.2%;	Score 815.6;	DB 22;	Length 1529;
Best Local Similarity	75.4%;	Pred. No. 1.8e-243;		
Matches 1126;	Conservative 4;	Mismatches 325;	Indels 39;	Gaps 8

OY	1	gacatgctcgaagaactaacgctgvcgvcgctctttaaactgcaatctgaacgvggtag	60
Db	8	gatactgctcgaatctgaaactgvcgvcgaactgcttaacaaatggaatctgaaacgaacg	67
OY	61	caatac-----ctagvcgcggaacggtgagtaacagctgtaacttcc	104
Db	68	atgctagctctgacgagcgctcgagctgvcgacggctgagtaatgatalagatctacc	127
OY	105	tcagagctctggataactcttcgaagaagaaagctaaccggatgctctgttgatcac	164
Db	128	taacagctgggggaacaacctgtgtgaaaacagacttaacatccgatactccaaacggcga	187
OY	165	aagatttatagtgaaagatttatgtcttggaagtggccgcgcgcgacttaactagt	224
Db	188	agca--gggaacctcggccttcgcgtataatagtgcctatgtctggattgctagtctg	245
OY	225	gtgaagtaatgctcacaaagcgcgaactcgtgacccgcctgaaagagctgctccgcac	284
Db	246	gtgaggtaaatgcccccaagcgaacgactccgtagctgtgttgaggtatgatacgaac	305
OY	285	aatgaaactggagacaacgttccatactcttaacgggaagcagagctaaagaatctgtcca	344
Db	306	actgggaactggaacacgcggccagactctcctaaagggaagcagaagtgggaatatgtgcaa	365
OY	345	tgggggaacaccttgagcagcgaacgcgcgtgaacggaaagaaagtctcggattgaaat	404
Db	366	tggaggaacctctgatactgacgaatagccgctgtgtgtgaaagagccttaaggtgtgaaagc	425
OY	405	tcaatgagcaggaaaaaataagacagcaatgtgatactgtaactgctca-----	452
Db	426	actctcaatgaaggaaaaatctlaaagttlaataacttaagccctgaactgacttaactcag	485
OY	453	---aagcacccgctaacactgtgccagaagccgcggtaatacgtga--tgggtgcaagcgttg	508
Db	486	aagaagccacccgcttaacactccgcgtgccagaagccgcggttaatacggaaagggctgcaagcgtta	545
OY	509	tctcgaaatcatctggcgtgaaagggctgctgaagcggattgtgaaatctacggctgtgaaactg	568
Db	546	atcgaatatctgvcgctgaaagccgcgttaagcggcttaaaacaagctcagatgtgaaagccc	605
OY	569	cgggctcaaccgcgtgctctgcaattgaaactacaacagctctggaagtgttgaggaggaacgttg	628
Db	606	cgggctcaaccgcgtggaactgcaattgaaacagcttgaactgaagctgtgtgtagaaggaagctg	665
OY	629	gaattccagcgttgaaacggtgaatctgctgatactctggaagaaacaacagctgvcgaagcgtg	688
Db	666	gaattccaagctgtagcggctgaaatgtcgatgatactctgaaggaacaacaacagctgvcgaagcgtg	725
OY	689	actgtcgtgctcaaaaactgacgcgtgagggcagcaaacgctgtggtgtgtgtaaacgggatactg	748
Db	726	gtctctcttgaccaacaacatgacgcgtgtgagctgtggaagacgctgtggtgtgacaacgggatactg	785
OY	749	accgcggtaatccacgcgccttaacgtgtctcaacagctgtgggggtgtttaaaccctgaat	808
Db	786	accocggtagtccaagccgcttaaacgatagttcaactcaactgctgtgggcgggttttcgcgtagt	845
OY	809	aacgaacctaaagataagtagaacgccttgvgggactagtctgcgaagatgaaactcaa	868
Db	846	gattga--ttaacgcataatgttaacccctctgggaatgatacgcgcgaagctttaaactcaa	904





Db	856	ccctttagtgctgcagcctaaagcattaaagcattccgccttcggtggagtaagcgtcgaagaactg	915
Qy	861	aaactcaaaagaaattgaacgagggttcgcgaacaacggttgagacatggtgtttaattcagtg	920
Db	916	aaactcaaaagaaattgaacgagggttcgcgaacaacggttgagacatggtgtttaattcagag	975
Qy	921	ataccccaataaacctcaaccctgggcttgacat- gtaacttgtaacatgtagaatalagag	979
Db	976	caacgcgaagaaaccttaacacaggtccttgacaatacctctgacaataccctaaagataagagatcc	1035
Qy	980	ccttc--gggcagattcaacaggtgctgcatagggttcgtcagctcgtgtcgttgagatgtt	1037
Db	1036	ccttcggggcgagatgacagagtggtgacatggttcgtcgcagctcgtgtcgttgagatgtt	1095
Qy	1038	gggtttaagtcctcgcaacagagcgaacccct- atcgtatgttgcataccttaagttgggac	1096
Db	1096	gggtttaagtcctcgcaacagagcgaaccccttga- tcttgctgcagaatcagttgggac	1155
Qy	1097	tgtgacgaatactcggttgacaacccggaagagcggtgagatgacgttcaatcctcagtg	1156
Db	1156	tctaaagtgactgcgcggtgacaaacccggaggaaggttggtgatagacgttcaatcactatgc	1215
Qy	1157	ccttattgcacagggccacaacacagtgctacaat- tggccgatatcagaggggtcgcacaactgcg	1216
Db	1216	cccttatgaccttggctacacacagtgctacaat- tgacagaaacaaaggcagcgaacccgc	1275
Qy	1217	aagggagagcttaactcctaanaagtcgtcccggttc- cgatgtgggttcgtcaactcgaccc	1276
Db	1276	gaagggttaagccaatacccaacaacctcgttcc- tccagtttcgtgactgcagttcgcgaactcgactg	1335
Qy	1277	catgaagtcggaatcgcctagtaatacgcggatcaga- catgacgtgcgggtgaaatacgttcccgga	1336
Db	1336	cgtagaactggaatcgtctagtaatacgcggatcaga- catgacgtgcgggtgaaatacgttcccgga	1395
Qy	1337	cctgtgacaacaacccgcgtgacaccccgagttgg- ggagacaccggaagtgtcttctgtta	1396
Db	1396	cctgtgacaacaacccgcgtgacaccccgagagtt- gttaaacaccggaagtgtcttctgtta	1455
Qy	1397	accgtaagaggaacagactactaatacgtgaatac- tcgttaaaagggtgataagtcgttaacaagt	1456
Db	1456	ccttttagggcgccagcgccgaaggtggggacagat- gatgttggttggaagtcgttaacaagt	1515
Qy	1457	acc 1459	
Db	1516	agc 1518	
RESULT			
6			
ID	AAK83569	standard; DNA; 1535 bp.	
XX	AAK83569;		
AC	AAK83569;		
XX	21-DEC-1999	(first entry)	
DT	16S rDNA gene fragment from marine bacterium, isolate K3-3.		
XX	Monitoring; oil; contamination; sea water; detection; flagellum;		
KW	Gram-negative bacterium; Proteobacteria; glucose; carbon source;		
KW	alkane; 16S rDNA gene; ds.		
KW	Proteobacteria.		
OS	JP11243967-A.		
XX	14-SEP-1999.		
PD	04-MAR-1998; 98JP-0069399.		
XX	04-MAR-1998; 98JP-0069399.		
XX	(KAIY-) KAIYO BIOTECHNOLOGY KENKYUSHO KK.		

DR WPI: 1999-564435/48.

XX

PT Monitoring of oil contamination of sea water - where oil contamination

PT is evaluated by detection of a microbe having properties from e.g.

PT having no flagellum, being a Gram-negative bacterium, belonging to

PT Proteobacteria, gamma subdivision, etc.2

XX

PS Claim 3; Page 7-8; 15pp; Japanese.

XX

CC The invention relates to a method for monitoring oil contamination of

CC sea water by detecting, in the sea water, a microbe having the following

CC properties: (1) it has no flagellum; (2) it is a Gram-negative bacterium;

CC (3) it belongs to Proteobacteria, gamma subdivision; (4) it cannot

CC assimilate glucose as a single carbon source; and (5) it efficiently

CC assimilates at least one of 10-30C n-alkanes. This sequence represents

CC a fragment of the 16S rDNA gene from the microbe of the invention,

CC isolate K3-3.

XX

XQ Sequence 1535 BP; 375 A; 350 C; 496 G; 314 T; 0 other;

Query Match	55.28;	Score 814.6;	DB 20;	Length 1535;
Best Local Similarity	76.08;	Pred. No. 3.7e-243;		
Matches 1137;	Conservative	0;	Mismatches 319;	Indels 41; Gaps 9

OY	1	gaatagtgctgaagcaactaaagcttgagcgcgctctcaaaatcacaatctgaagcgaggtag	60
Db	9	gattccgtgctcaagatctgaacgcgtgagcgccagcctcaacacatactcaatctgagcgaaacg	68
OY	61	caatcac-----ctagcgcggaacgggtgagtaacagctgtaattctcc	104
Db	69	atgttagcttgctacacagcgctcgagcgcggaagcggttgagtaacgctggaattctgcc	128
OY	105	tcgagctctgagataaacttccggaaagaaagcctaataccggatagtcctgtgagtcac	164
Db	129	catttgtagggataaactctgggaaactcaagcctaataccgataacccctacgggggaa	188
OY	165	aagaatttgatagtaagaagatttaactcttgagatatgagcccgcgccgaattagctatg	224
Db	189	agca--gggagccttcggtctgtgagatgtgagtcgcgcgtcgattagctagttg	246
OY	225	gtgaggtgtaatgcttcacacaaagcgagcatctggtagccggccttagagaggtgtccggccac	284
Db	247	gtgggtgtaaaggtccctacaaagcgagcatccgtgtagctgtgtagagtgatcaacgacac	306
OY	285	aatggaactgtgagacacggtccatgctctctacggaggaagcagaagcttaagaatctgtgca	344
Db	307	acgggagactgagaaacagcgccggagctctcctaaggaagcagcagtgtaggaattcttgaca	366
OY	345	tgggtgaaaccctgtgaagcagcagcgccggtgaaagagaaggtctctggattgttaagt	404
Db	367	tgggtgcaaccctatcatcagcgcattgctgtgtagaagaagcgcttagagttgttaaacg	426
OY	405	tcattagagcaggaanaata-----agcagcaatgtgatgtatgataccggcct	451
Db	427	aatttcagcagagggaggaagccttcgatttaatactctgggtactgtagactctaccggac	486
OY	452	a--aagcaacggctaactacgtgcccagcagcgcggttaatacgtatgtgcaagcgttgt	509
Db	487	aagaagcacccgcttaattctgtgccacacagcgcgtgtaatacgaagaagttgcggacttaa	546
OY	510	tcggaatcatgtggcgctaaagaaggtgctgtagcggaattgttaadtcaagcttgaaacctgc	569
Db	547	tcggaatcaacgtggcgctaaagcgcgtgtagcggttctgttaagtcaaaatgtgaaagcccc	606
OY	570	gggtctcaaacccgtgctgtgacgtctgaaactcaaaagctctgtagtcttgtagaagcgaatg	629
Db	607	gggtcacaaccgtggaactgctattgtgaacttgcagcttagagtgacgttagagggaggtg	666
OY	630	aattccaggtgttaacggttgaatgtcgtatgatactctgtagaagaacaacagctgtagaagcga	689
Db	667	aattccggtgttaacggttgaatgtcgtatgagatactgtagaagaacaacccgtgtagaagcga	726

OY	690	cttgcgtgcctcaaaccttgcgtctgagcgacgcaagaagcgttgggttgatgttaacgggattatgata	749
OY	727	cctcttcgtgccttgacacattgacgtctggagggttgcgcgaagcgttggggagcaaacggattatgata	786
OY	750	ccctgggttaatccacgcgccttaaacgtctgtctacccaggttctgttgggggtttta--acccctcag	807
OY	787	cccttgattatccacgcgcgttaaacggatgtctactactgttcgtttgggacttagttatctcttgg	846
OY	808	taacgaacctaacgcgattgaagtatagacgcgcctctggggacttatgtcttcgcaagggtgaacctca	867
Db	847	tgcgcgaagtataacgcgatatgaatagtagacgcgcctctggggagtagtcggccgaaggttaaacctca	906
OY	868	aaggaattatcacgggggtctccgcacaagcggctggagacatgtgttttaattcgtatatacccc	927
Db	907	aatgaattatgcaggggggtcccgccacaagcgggtggagacatgtgttttaattcgtatatacccc	966
OY	928	aaaaaacctcaaccttggcgtcttgacatgatactga-atcatgtatagatatatagaacctctgg	986
Db	967	aagacaccttcaacgcgcctcttgacatcttcggaaaccttctcagaagatatagatttgggtcctcgg	1026
OY	987	g--cagaattacacgggtgcgcgcattgtttgttcgtcaacatctgtctctggagatgttgggttaa	1044
Db	1027	gaacgcagtgacacgggtgcgcgcattgtttgttcgtcaacatctgtctctggagatgttgggttaa	1086
OY	1045	gtcccgcaagaagcgcgaacacctatcgtatgtttgcta--cccttaagtttgggcactgtgtac	1102
Db	1087	gtcccgtaagaagcgcgaacaccttgcctcttagtttgcacaacattcgggttgggaaacctcaga	1146
OY	1103	gaacatcgcgggtgacaaacccggaggaagcgcgggattgacgttcaaatctctatctgtccctta	1162
Db	1147	gagactcgcgggtgacaaacccggaggaagcgcgggattgacgttcaaatctctatctgtccctta	1206
OY	1163	tgtccagggcgacaacacgttgtctacaaatgtgcgcatacagaagggtctgcacaactcgcgaagag	1222
Db	1207	cgcgccttgctacacacacgttgtctacaaatgtgcgcatacagaagggtcgaagtcgcgaagcc	1266
OY	1223	gagcttatctctaaagttcgttcccgattccgatttgggttgggtctgcgaacctcgcgccccatgaa	1282
Db	1267	aagcaaatccctttaaaccgtctgttgcgtccggatttggagatcgcgaacctcgcgccccatgaa	1326
OY	1283	gtcgcgatacgtctaaatctgcgcgatacagaatgcgcgcgtgtgaatacgttctccggagacctgt	1342
Db	1327	gtcgcgatacgtctaaatctgcgcgatacagaatgcgtgcgtgtgaatacgttctccggagacctgt	1386
OY	1343	aaacacgcgcgcgtacacacacactgagttggggagcaaccggaaatgtgtctttgttaacgta	1402
Db	1387	aaacacgcgcgcgtacacacacactgagttggggagcaaccggaaatgtgtctttgttaacgta	1445
OY	1403	aggagacagactataaggttggaacctgtataaagggggttgaatctgtataacaaagtatcc	1459
Db	1446	ggagagacagactataacacagcgtgtgtgttcatgtacgcggggtgaagtctgtataacaaagtatcc	1502
RESULT 7			
ID	AAC86030 standard; CDNA. 1485 BP.		
XX	AAC86030;		
XX	29-AUG-2001 (first entry)		
XX	16s rDNA, Probel.		
XX	16S rDNA; polyphosphate accumulating organism: PAO; probe: primer;		
KW	detection: phosphorus; waste water; sludge; ss..		
XX	Propionibacter pelophilus.		
PN	WO200146459-A1.		
PD	28-JUN-2001.		
PF	28-DEC-2000; 2000WO-AU01611.		

```

PR    23-DEC-1999;      99AU-0004867.
XX
PA    (CRCW-) CRC WASTE MANAGEMENT & POLLUTION CONTROL.
XX
PI    Hugenholz P, Crocetti GR, Tyson GW, Blackall LL:
DR
XX
XX    MPI: 2001-408656/43.
XX
PT    Novel oligonucleotide probe or primer useful for detecting
XX    polyphosphate accumulating organism in a sample, comprises a sequence
XX    that is unique to 16S rDNA of polyphosphate accumulating organisms -
PS
XX
XX    Claim 4; Fig 3; 54pp: English.
XX
CC    The sequences given in AAC86021-30 represent 16S rDNA sequences from
CC    polyphosphate accumulating organisms (PAOs). Sequences which are unique
CC    to these 16S rDNA sequences are used to create a probe or primer for
CC    detecting the relevant organisms. The primer/probe sequences are
CC    useful for detecting PAO cells in a sample, by treating cells in the
CC    sample to fix cellular contents, contacting fixed cells with the
CC    primer/probe which is labelled with a radiolabel, a reporter group or a
CC    hapten, under conditions which allow the probe to hybridize with 16S
CC    rRNA within the fixed cell, removing unhybridized probe from the fixed
CC    cells, and detecting the labeled probe-RNA hybrid by fluorescence in
CC    situ hybridization. The primer/probe sequences are useful for
CC    identifying PAOs that are capable of biologically removing
CC    phosphorus from waste water. Rapid assessment of the presence of a
CC    number of PAOs in a waste water sample, can be done using the primer/
CC    probe sequences. They allow quick and convenient assessment of whether
CC    a sludge or waste water sample includes PAOs and allows quantitation
CC    of PAO cells in samples.
XX
XX    Sequence 1485 BP; 367 A; 341 C; 482 G; 295 T; 0 other:
SQ
Query Match          55.1%; Score 814.4; DB 22; Length 1485;
Best Local Similarity 75.7%; Pred. No.4.2e-243;
Matches 1127; Conservative 0; Mismatches 321; Indels 40; Gaps
QY       7 ggcctagaactaacgctcgccggcgcttctaataatgcgaagtcaggcgg----- 55
DB        1 ggtctagattgcaacgctcgccggcgcatctcttaaccacgaatcgaaagcgacgatggcg 60
QY       56 ----ggttagcaatccctagtcggcgcaaggggttgagttaacacgvtgtaactcttcctccagtc 112
DB        61 ctgtcaccttgatggcggaugtggcgaaaggggtgtagtaatgatcatcggaacgfatccccggagtc 120
QY       113 tgggataaactttccggaaggaagaaagctaataccggatagcttcgttggatcacacaagattcg 172
DB       121 gggtataaagctagtcgaaagttaaccttaataccgcatattcttgagcagagaagaaggggg 180
QY       173 ataggtaaagatttatctgtcttgagagatgagccgggccgagatagctagtgttgtaggta 232
DB       181 atcg--caagacctcttcgcttctcgagcgcgacgcatgctgattagctaagtctggtagggta 238
QY       233 atggcttcaaccaagcgacgacgttcgtagccgctcgtagagggvtgtccggccaacaatggnaac 292
DB       239 aaagctcataccaagcgacgacgttcgtagcggtcttgtagagagatgatccgccacacttggac 298
QY       293 ttgagacacggtlccaatactctactacggaaggaagcagcaatgaagaatctgtcctaattgggggaa 352
DB       299 ttgagacacggtcccgactctactacggaaggaagcagcttgggaatatcttggacaattgggcga 358
QY       353 accctgaagcagcgacgcccgttgaaacgaagaaggtcttcggatttgaagtcatatagg 412
DB       359 agccttgatccaagcatalgvcggtgagtgaaagaaagccttcgggttgtaaagccttctcgg 418
QY       413 caggaaaaaatgaagc-----agcaatgtgatatgtgtacacctgcta--aagca 457
DB       419 tcgggaagaagaattgcaacgctctaatacatagcgtgtgtatgtacgtaaccgacataagaagca 478
QY       458 ccggactaactacgtgcgaacgacccgggttaataagtatgtltgcaagagctgtgttcggatc 517

```





```

Db      1448 ccttattgagccagccgcgaagctgaggacagatgattggtgtgaaagtcgtaacaaggt 1507
QY      1457 acc 1459
Db      1508 aac 1510

RESULT 9
AAK83564
ID      AAK83564 standard; DNA; 1529 BP.
XX
AC      AAK83564;
XX
DT      21-DEC-1999 (first entry)
XX
DE      16S rDNA gene fragment from marine bacterium, isolate SF-T1.
XX
KW      Monitoring; oil; contamination; sea water; detection; flagellum;
KW      Gram-negative bacterium; Proteobacteria; glucose; carbon source;
KW      alkane; 16S rDNA gene; ds.
XX
OS      Proteobacteria.
XX
PN      JF11243967-A.
XX
PD      14-SEP-1999.
XX
PF      04-MAR-1998; 98JP-0069399.
XX
PR      04-MAR-1998; 98JP-0069399.
XX
PA      (KAIY-) KAIYO BIOTECHNOLOGY KENKYUSHO KK.
XX
DR      WPI; 1999-564435/48.
XX
PT      Monitoring of oil contamination of sea water - where oil contamination
PT      is evaluated by detection of a microbe having properties from e.g.
PT      having no flagellum, being a Gram-negative bacterium, belonging to
PT      Proteobacteria, gamma subdivision, etc.z
XX
PS      Claim 3; Page 5; 15pp; Japanese.
XX
CC      The invention relates to a method for monitoring oil contamination of
CC      sea water by detecting, in the sea water, a microbe having the following
CC      properties: (1) it has no flagellum; (2) it is a Gram-negative bacterium;
CC      (3) it belongs to Proteobacteria, gamma subdivision; (4) it cannot
CC      assimilate glucose as a single carbon source; and (5) it efficiently
CC      assimilates at least one of 10-30C n-alkanes. This sequence represents
CC      a fragment of the 16S rDNA gene from the microbe of the invention,
CC      isolate SF-T1.
XX
SO      Sequence 1529 BP; 379 A; 351 C; 487 G; 312 T; 0 other;

Query Match          54.9%; Score 811.6; DB 20; Length 1529;
Best Local Similarity 76.2%; Pred. No. 3.1e-242;
Matches 1142; Conservative 0; Mismatches 314; Indels 42; Gaps 10;

```

```

QY      1 gatcatgctcagaactaacgtctgctgcgcgtcttaacatgcaagtcgagcggtgag 60
Db      2 gatcctgctcagattggaacgctgctgcgcgcgtcttaacatgcaagtcgagcggtgag 61
QY      61 caatc-----ctagcggcggaagcggtgagtaacagctgtgtaactcc 104
Db      62 atcctagcttctagagagcgctcgagcggtgagtaacagctgtgtaactcc 121
QY      105 tcgagctcggtgataacttccgaaagaaagctaataccggtatgctctgttgatacac 164
Db      122 cattgagaggggataaccctggggaaccaggtctaataccggtatgctcttgatacac 181
QY      165 aagatttgatagtaagaattatgcttggatgagatgagccggtcgatgactgaagtg 224
Db      182 agcaggggat--ctcgaccttgcgtgatgatgagatgagctgcgtcgatgactgaagtg 239

```

```

OY 225 gtcgaagtaatggtcaccacgaagcgacgacgagtgagcgagcgttcgagaggggtgcgcgcac 284
    |||||
DB 240 gtagagtaatggtcaccacgaagcgacgacgagtgagcgagcgttcgagaggggtgcgcgcac 299
OY 285 aatggaactgagacacggttcacacgagcgttcgagagggcgagcgttcgagaggggtgcgcgcac 344
    |||||
DB 300 accggaactgagacacggttcacacgagcgttcgagagggcgagcgttcgagaggggtgcgcgcac 359
OY 345 tgggggaacccctgagacgagcgttcgagagggcgagcgttcgagaggggtgcgcgcac 403
    |||||
DB 360 tgggggaacccctgagacgagcgttcgagagggcgagcgttcgagaggggtgcgcgcac 419
OY 404 ---ttcatatgacgagaaatgagcaatgtg-----atgagtgacacgtcct 451
    |||||
DB 420 actttcagtagaggaagagcgttcacacgagcgttcgagagggcgagcgttcgagaggggtgcgcgcac 479
OY 452 a--aagcaccggtcactacggttcgagacgagcgttcgagagggcgagcgttcgagaggggtgcgcgcac 509
    |||||
DB 480 aagaaacacggtcactacggttcgagacgagcgttcgagagggcgagcgttcgagaggggtgcgcgcac 539
OY 510 tgggaatcatgtgggtcgaagagtgagcgttcgagagggcgagcgttcgagaggggtgcgcgcac 569
    |||||
DB 540 tgggaatcatgtgggtcgaagagtgagcgttcgagagggcgagcgttcgagaggggtgcgcgcac 599
OY 570 gggcgaacccgttcgagacgagcgttcgagagggcgagcgttcgagaggggtgcgcgcac 629
    |||||
DB 600 gggcgaacccgttcgagacgagcgttcgagagggcgagcgttcgagaggggtgcgcgcac 659
OY 630 aatccgaagtgcagcgttcgagagggcgagcgttcgagaggggtgcgcgcac 689
    |||||
DB 660 aatccgaagtgcagcgttcgagagggcgagcgttcgagaggggtgcgcgcac 719
OY 690 cttgctggtcgaagcgttcgagagggcgagcgttcgagaggggtgcgcgcac 749
    |||||
DB 720 cttgctggtcgaagcgttcgagagggcgagcgttcgagaggggtgcgcgcac 779
OY 750 ccccggaacccgttcgagacgagcgttcgagagggcgagcgttcgagaggggtgcgcgcac 807
    |||||
DB 780 ccccggaacccgttcgagacgagcgttcgagagggcgagcgttcgagaggggtgcgcgcac 839
OY 808 taacgaactacgagcgttcgagagggcgagcgttcgagaggggtgcgcgcac 867
    |||||
DB 840 taacgaactacgagcgttcgagagggcgagcgttcgagaggggtgcgcgcac 899
OY 868 aaggaatgcagcgttcgagagggcgagcgttcgagaggggtgcgcgcac 927
    |||||
DB 900 aaggaatgcagcgttcgagagggcgagcgttcgagaggggtgcgcgcac 959
OY 928 aaaaactcaccgttcgagacgagcgttcgagagggcgagcgttcgagaggggtgcgcgcac 986
    |||||
DB 960 aaaaactcaccgttcgagacgagcgttcgagagggcgagcgttcgagaggggtgcgcgcac 1019
OY 987 --gcagatcacagcgttcgagacgagcgttcgagagggcgagcgttcgagaggggtgcgcgcac 1044
    |||||
DB 1020 gcagatcacagcgttcgagacgagcgttcgagagggcgagcgttcgagaggggtgcgcgcac 1079
OY 1045 gtcggaacccgttcgagacgagcgttcgagagggcgagcgttcgagaggggtgcgcgcac 1101
    |||||
DB 1080 gtcggaacccgttcgagacgagcgttcgagagggcgagcgttcgagaggggtgcgcgcac 1139
OY 1102 cgaacatgcggttcgagacgagcgttcgagagggcgagcgttcgagaggggtgcgcgcac 1161
    |||||
DB 1140 ggaagctgcggttcgagacgagcgttcgagagggcgagcgttcgagaggggtgcgcgcac 1199
OY 1162 atgtccagggcgagcgttcgagacgagcgttcgagagggcgagcgttcgagaggggtgcgcgcac 1221
    |||||
DB 1200 acggaacccgttcgagacgagcgttcgagagggcgagcgttcgagaggggtgcgcgcac 1259
OY 1222 ggaagcgaactcgaagcgttcgagacgagcgttcgagagggcgagcgttcgagaggggtgcgcgcac 1281
    |||||
DB 1260 caagcgaactcgaagcgttcgagacgagcgttcgagagggcgagcgttcgagaggggtgcgcgcac 1319

```

```

OY 1282 agtcgaatcgtcagtagtaacgagacgagcgttcgagagggcgagcgttcgagaggggtgcgcgcac 1341
    |||||
DB 1320 agtcgaatcgtcagtagtaacgagacgagcgttcgagagggcgagcgttcgagaggggtgcgcgcac 1379
OY 1342 tacacacccggttcgagacgagcgttcgagagggcgagcgttcgagaggggtgcgcgcac 1401
    |||||
DB 1380 tacacacccggttcgagacgagcgttcgagagggcgagcgttcgagaggggtgcgcgcac 1438
OY 1402 aaggaacgagcgttcgagacgagcgttcgagagggcgagcgttcgagaggggtgcgcgcac 1459
    |||||
DB 1439 cgggaagcgttcgagacgagcgttcgagagggcgagcgttcgagaggggtgcgcgcac 1496

```

RESULT 10  
AA231557  
ID AA231557 standard; rRNA; 1508 BP.

AA231557;  
13-JAN-2000 (first entry)

DE Bacillus pectate lyase rRNA sequence.

KW pectate lyase; polysaccharide lyase; enzyme; pectin degradation;  
KW polygalacturonide; detergent composition; hard surface treatment;  
KW cellulosic fibre; plant material degradation; recycled waste paper;  
KW mechanical paper-making pulp; wine processing; ss.

OS Bacillus sp.

PN WO9927084-A1.

PD 03-JUN-1999.

PF 24-NOV-1998; 98WO-DK00515.

PR 24-NOV-1997; 97DK-0001343.

PR 06-MAY-1998; 98US-0073684.

PR 02-NOV-1998; 98US-0184217.

PA (NOVO ) NOVO-NORDISK AS.

PI Andersen LN, Schuelein M, Lange NEK, Bjornvad ME, Moller S;

PI Glad SOS, Kauphagen MS, Schnorr K, Kongsbaek L;

DR WPT: 1999-610579/52.

DR P-PSDB: AAY43215.

PS New isolated pectate lyase enzymes -

PS Claim 8: Page 109-111; 113pp; English.

This sequence encodes a *Bacillus* species pectate lyase of the  
CC invention. The pectate lyase enzymes are obtained from *Bacillus*  
CC agaradhaerens (BA), *Bacillus* licheniformis (BL), *Bacillus* halodurans (BH)  
CC and other *Bacillus* species. The pectate lyase enzymes can be used for  
CC degrading pectin, pectate and polygalacturonides. They can be used in  
CC detergent compositions, for cleaning hard surfaces, for machine treatment  
CC of fabrics, for improving the properties of cellulosic fibres, yarn,  
CC woven or non-woven fabric, for the degradation of plant material (e.g.  
CC recycled waste paper, mechanical paper-making pulps or fibres subjected  
CC to a retting process, for preparing animal feed or for processing wine  
CC or juice. DNA encoding the enzymes can also be used for the production of  
CC transgenic plants.

SQ Sequence 1508 BP; 384 A; 348 C; 464 G; 312 U; 0 other;

Query Match 54.9%; Score 810.2; DB 20; Length 1508;  
Best Local Similarity 61.7%; Pred. No. 8; Se-242;  
Matches 921; Conservative 220; Mismatches 303; Indels 49; Gaps 10;





```

XX Monitoring; oil; contamination; sea water; detection; flagellum;
KM Gram-negative bacterium; Proteobacteria; glucose; carbon source;
KW alkane; 16S rDNA gene; ds.
XX
XX Proteobacteria.
XX
XX JP11243967-A.
XX
XX 14-SEP-1999.
XX
XX 04-MAR-1998; 98BP-0069399.
XX
XX 04-MAR-1998; 98BP-0069399.
XX
XX (KAIY-) KAIYO BIOTECHNOLOGY KENKYUSHO KK.
XX
XX WPI; 1999-564435/48.
XX
XX Monitoring of oil contamination of sea water - where oil contamination
XX is evaluated by detection of a microbe having properties from e.g.
XX having no flagellum, being a Gram-negative bacterium, belonging to
XX Proteobacteria, gamma subdivision, etc.z
XX
XX Claim 3; Page 6-7; 15pp; Japanese.
XX
XX The invention relates to a method for monitoring oil contamination of
XX sea water by detecting, in the sea water, a microbe having the following
XX properties: (1) it has no flagellum; (2) it is a Gram-negative bacterium;
XX (3) it belongs to Proteobacteria, gamma subdivision; (4) it cannot
XX assimilate glucose as a single carbon source; and (5) it efficiently
XX assimilates at least one of 10-30C n-alkanes. This sequence represents
XX a fragment of the 16S rDNA gene from the microbe of the invention.
XX Isolate Sim-2.
XX
XX Sequence 1529 BP; 377 A; 352 C; 488 G; 312 T; 0 other:
XX
XX
XX Query Match 54.8%; Score 810; DB 20; Length 1529;
XX Best Local Similarity 76.2%; Pred. No. 9,8e-242;
XX Matches 1141; Conservative 0; Mismatches 315; Indels 42; Gaps 10;
XX
XX 1 gatcatgctcagaactaacgtctgagcgagctttaaactgcaagtcgagcggtag 60
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 2 gatctgctcagaactggaacgtgagcgagcctaacaacatgcaagtcgagcgaaag 61
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 61 caatc-----ctagcgcggaacggtgagtaaacgtgtgaactcc 104
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 62 atcctagcttctagagagcgctcagcgagcgagcggtgagtaaacgtgtgaactcc 121
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 105 tccgagctcagataacttccgaagaagaactaataccgagatgctctgttgataac 164
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 122 ccttgagtgaggaatacccggggaacccagctataaccgataatccctcaggggaa 181
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 165 aagatttgataggaattatgcttgagagatgagccgagcgagctgagctgagctg 224
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 182 agca--ggggaacctcgagcctgtgtgagatgagctcgagctgagctgagctg 239
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 225 gtagagtaatgctcacaagcgagacatcggtagccgagagagatgttcgagcaac 284
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 240 gtagagtaatgctcacaagcgagacatcggtagctggtcttaagagagtgatcagcaac 299
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 285 aatggaactgaacacggtcactactactcagggagcgagctgtaagaactctgtcaca 344
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 300 accgggaactgagacacgagccggaactcctcaggaagcgagctggaactcttgaca 359
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 345 tgggggaaacctgaagcgagacgagcggtgaaacgaagaagctctgagattgtaag- 403
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 360 tgggggaaacctgagccagcctgagcggtgtgtgagaaagagccttcgggttgtaagc 419
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 404 ---tcaatgagcagaaataacagcaatgtg-----atgacgttaactgct 451
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 420 acttcaagtaggaggaaggctatacttaatacgaatgagtaactggaagcttaaccag 479

```

```

OY 452 a--aagcaccggttaactactgtccagcagccggttaactgtaatggtcgaagcgtgt 509
OY | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 480 aagaagcaccggttaacttctgctcagcagccggtgtaactgaagaagtgcaagcgttaa 539
OY | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 510 tgggaactatggcggtlaaagggtgctagcgagattgttaagtcaggtgtgaaactgtc 569
OY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 540 tgggaactatggcggtlaaaggcgtgtagcggttctgtaagtcagatgtgaaagcccc 599
OY | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 570 gggcttaacccgtgctgacgttgaacttgaactgaactgtgagtttgaggagagcaggtg 629
OY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 600 gggcttaacccgtggaactgtgacgttgaacttgaacttgaactgaactgaactgaactgtg 659
OY | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 630 aatcgaagtgtagcggtgaactgtcagatctgtagaggaacacacagctgaggaagcgca 689
OY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 660 aatcgcgtgtgtagcggtgaactgtcagatctgtagaggaacacacagctgaggaagcgcg 719
OY | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 690 ctgtgctgctcaaaactgacgtctgagcgacgaaagcgtgtgtgtagtaacgagattagata 749
OY | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 720 cctcctgagctgacatgacgtctgagcgaggtgtggaagcgtgtgtggaaggaattagata 779
OY | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 750 ccccggaatccagcccttaaacgttctgtaaccagctgtgtgtgtgtta--accctcag 807
OY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 780 cctgtgtagtccagccgtaaacgagatgctcagctagctgtgtgtgtgtgtgtgtgtgtgt 839
OY | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 808 taacgaacttaacgagattagtagacgcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 867
OY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 840 tgcaggaagtttaacgagtaagtagacgcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 899
OY | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 868 aaggaatgacgaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 927
OY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 900 atgtaattgacggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 959
OY | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 928 aaaaactcaactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 966
OY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 960 aagaacttaacgagcttgaacttgaacttgaacttgaacttgaacttgaacttgaacttgaact 1019
OY | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 987 --gagattcaacgagtgctgacatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1044
OY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1020 gagcaagtgtagaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1079
OY | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1045 gtccgcgaacgagcgcaacccctactgtatgtgc--taactaaattgtggcaactgtga 1101
OY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1080 gtccgcgaacgagcgcaacccctactgtatgtgc--taactaaattgtggcaactgtga 1139
OY | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1102 cgaactgtcggtgacaaacccggaagcggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1161
OY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1140 gtagactgtcggtgacaaacccggaagcggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1199
OY | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1162 atgtccgaagcgcaacgagtgctacatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1221
OY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1200 acgagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1259
OY | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1222 gtagactgtcggtgacaaacccggaagcggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1281
OY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1260 caaggaatactccttaaaactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1319
OY | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1282 agtcggaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1341
OY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1320 agtcggaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1379
OY | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1342 tacacacgcccgtacacacacactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1401
OY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1380 tacacacgcccgtacacacacactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1438
OY | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 1402 aaggaagacacttaaggtgtaaacctgttaaaaggggtgtgaagtcgttaaaaggttacc 1459
OY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1439 cgggagagacgttaccacacaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1496

```

RESULT 13  
 AAX83565  
 ID AAX83565 standard; DNA; 1529 BP.



```
XX AAX83565;
XX
XX 21-DEC-1999 (first entry)
XX
XX 16S rDNA gene fragment from marine bacterium isolate Wf-1.
XX
XX Monitoring; oil; contamination: sea water; detection: flagellum;
XX Gram-negative bacterium; Proteobacteria; glucose; carbon source;
XX alkane; 16S rDNA gene; ds.
XX
XX Proteobacteria.
XX JP11243967-A.
XX
XX 14-SEP-1999.
XX
XX 04-MAR-1998; 98JP-0069399.
XX
XX 04-MAR-1998; 98JP-0069399.
XX
XX (KAIY-) KAIYO BIOTECHNOLOGY KENKUSHO KK.
XX
XX WPI: 1999-564435/48.
XX
XX Monitoring of oil contamination of sea water - where oil contamination
XX is evaluated by detection of a microbe having properties from e.g.
XX having no flagellum, being a Gram-negative bacterium, belonging to
XX Proteobacteria, gamma subdivision, etc.2
XX
XX Claim 3: Page 5: 15pp; Japanese.
XX
XX The invention relates to a method for monitoring oil contamination of
XX sea water by detecting, in the sea water, a microbe having the following
XX properties: (1) it has no flagellum; (2) it is a Gram-negative bacterium;
XX (3) it belongs to Proteobacteria, gamma subdivision; (4) it cannot
XX assimilate glucose as a single carbon source; and (5) it efficiently
XX assimilates at least one of 10-30C n-alkanes. This sequence represents
XX a fragment of the 16S rDNA gene from the microbe of the invention,
XX isolate Wf-1.
XX
XX Sequence 1529 BP: 376 A; 352 C; 489 G; 312 T; 0 other;
SQ
Query Match 54.7%; Score 808.4; DB 20; Length 1529;
Best Local Similarity 76.1%; Pred. No. 3,1e-241;
Matches 1140; Conservative 0; Mismatches 316; Indels 42; Gaps 10;
QY 1 gatcatgctcagaactaactcgtgctgctgtcttaacatgacgtcagcggtag 60
DB 2 gatccgtgctcagatgacgtcgtgctgctcctaacaatgacgtcagcggtag 61
QY 61 caatac-----ctagcgcgaaacgggtgagtaaacgtgttaattcc 104
DB 62 atcttagctgtcaggagcgctgcagcgcgacgggtgagtaaacgtgttaattcc 121
QY 105 tccgaatcgtgataacttccgaagaagaacttaactcagatagtcctgtgtgacac 164
DB 122 cattagtggggataaacctcgggaaacccaagcctaatacgcatacctcctacggggaa 181
QY 165 aagatttgaatgaagaattatctgtcgtgagatgagccgcgagcgtatgaactgtg 224
DB 182 agca--ggggactctcggcctgtgtcgtatgtagtgacgtcgtcgtcttagctgtg 229
QY 225 gtagagtaatgctcaccgaagcgacgtatcgatcggcgtcgaaggggtgtccggcac 284
DB 240 gtagagtaatgctcaccgaagcgacgtatcgatcggcgtcgaaggggtgtccggcac 299
QY 285 aatggaacgagacacgtccatcaactcctacaggaagcagcttaagcttctgctca 344
DB 300 accggagactgagacacacgtccgagcctcctacaggaagcagcttggggatctctgacaa 359
QY 345 tgggggaaacccctgaagcagcgacgcgcgtgaacgaagaaggtcttcgattgtaaag- 403
```

```
DB 360 tgggggaaacccctgaagcagcgacgtcgtgtgtgaagaagcctcgtctgtgaagc 419
QY 404 ----ttcatgagcaggaataaataagcaatgtg-----atgagttaacctgct 451
DB 420 acttctagtaggaggaaggctatctccttaactagatagtaacttgaacttaccacag 479
QY 452 a--aaagcaccgcttaactcgtgcagcagcggtlaactagatgtgtcaagcgtgt 509
DB 480 aagaagcaccgcttaactcgtgcagcagcggtlaactagatgtgtcaagcgtgt 539
QY 510 tgggaatcattggtcgtaaaaggtgtcgtgagcgtatgttaagtcaggtgtgaatactgc 569
DB 540 tgggaatcattggtcgtaaaaggtgtcgtgagcgtatgttaagtcaggtgtgaatactgc 599
QY 570 gggctcaaccgtgtgcctcgtcacttgaactaactcgtgaggttctgtggaagcaggtg 629
DB 600 gggctcaaccgtgtgagcttgaacttgaactcgtgaggttctgtggaagcaggtg 659
QY 630 aattccaggtgtgagcgtgaaatgcgtatctgtggaaggaacacacaggtggtgaagcga 689
DB 660 aattccaggtgtgagcgtgaaatgcgtatctgtggaaggaacacacaggtggtgaagcga 719
QY 690 ctgtcgtgtcacaactgacgtcgtgagcagcgaagcgtgtgtgagtaaacggattagata 749
DB 720 cctcctgagactgacatgtgacgtcgtgaggtgtggaagcgtgtgtggaagcagcagattagata 779
QY 750 ccccggttaactcagcgccttaaacgtgtgtcactcaggtgtgtgtgtttaa--accctag 807
DB 780 cctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 839
QY 808 taacgacactaactgagttagtagacgtcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 867
DB 840 tgaaggaactaactgagttagtagacgtcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 899
QY 868 aaggaatctgacgggtgtccgcacaaagcgtgtgagcagctgtgtgtgtgtgtgtgtgtgtgt 927
DB 900 aatgaatctgacgggtgtccgcacaaagcgtgtgagcagctgtgtgtgtgtgtgtgtgtgtgt 959
QY 928 aaaaactcactcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 986
DB 960 aagaactcactcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1019
QY 987 --gcaattcacaagctgtcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1044
DB 1020 gggcgaagtgacacgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1079
QY 1045 gtcccgcaacgagcgaacccctatcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1101
DB 1080 gtcccgtaacgagcgaacccctatcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1139
QY 1102 cgaactcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1161
DB 1140 gtagactcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1199
QY 1162 atgtccagtgcaacacgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1221
DB 1200 acggcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1259
QY 1222 gtagactcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1281
DB 1260 caagcaaatcctccttaaaactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1319
QY 1282 agtcggaatcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1341
DB 1320 agtcggaatcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1379
QY 1342 tacacacgcgcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1401
DB 1380 tacacacgcgcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1438
QY 1402 aaggaagacgactaactgaagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1459
```







Mon Dec 17 07:48:10 2001

us-09-380-826a-1.rng

Page 18

QY	1215	gcacagagagacctaactctctaaagtcggtccagttcgttctggatcttggtgctctgcaactcagac	1274
Db	1268	gcgagggcgagagccaactctctaaacccgcttccagttctcgtatcttgaaagcgcgaactcgcc	1327
QY	1275	cccatgagacgcggaatcgcctagatcaatccgcggatccagcatgcgcggcgcggaatactcccg	1334
Db	1328	tacatgagcctgtgaatcgcctagatcaatccgcggatccagcatgcgcggcgcggaatactcccg	1387
QY	1335	gaccttgaacacccgcgcctgcacaccacttagtgaggagcaccgcgaagtgcctcttgt	1394
Db	1388	ggcctctgaacacacgcgcctgcacaccagagaatttctgtaacacccgcgaagtgcgtggggt	1447
QY	1395	taaccgtaagagagacagactactaaggtgaaactcgttaagggggttgaaatctgttaacag	1454
Db	1448	aaccttttgtaggcacagccgcctaaagctgtagacagatgatctgggggtgaagctcgttaacag	1507
QY	1455	gtacc	1459
Db	1508	gttagc	1512

Search completed: December 15, 2001, 03:29:41  
Job time: 7881 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 01:16:45 : Search time 172.39 Seconds  
(without alignments)  
1940.414 Million cell updates/sec

Title: US-09-380-826A-1

Perfect score: 1477

Sequence: 1 gacatcgccgcagactaac.....ccgtaatcgatccctgcag 1477

Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PCUTS.COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	814.2	55.1	1536	2	US-08-642-229A-1 Sequence 1, Appli
2	810.2	54.9	1508	4	US-09-198-955A-14 Sequence 14, Appl
3	805.4	54.5	1521	3	US-08-501-126-20 Sequence 20, Appl
4	803.8	54.4	1536	1	US-08-114-695A-7 Sequence 7, Appli
5	801.8	54.3	1512	3	US-08-995-960-2 Sequence 2, Appli
6	801.6	54.3	1555	2	US-08-757-653-160 Sequence 160, App
7	796.4	53.9	1506	4	US-09-198-955A-13 Sequence 13, Appl
8	792.4	53.6	1517	3	US-09-248-528-2 Sequence 2, Appli
9	792.4	53.6	1517	4	US-09-549-108-2 Sequence 2, Appli
10	792.4	53.6	1517	4	US-09-549-111-2 Sequence 2, Appli
11	792.4	53.6	1517	4	US-09-549-106-2 Sequence 2, Appli
12	792.4	53.6	1517	4	US-09-550-394-2 Sequence 2, Appli
13	790.6	53.5	1500	4	US-09-193-377B-3 Sequence 3, Appli
14	789.8	53.5	1516	3	US-09-248-528-3 Sequence 3, Appli
15	789.8	53.5	1516	4	US-09-549-108-3 Sequence 3, Appli
16	789.8	53.5	1516	4	US-09-549-111-3 Sequence 3, Appli
17	789.8	53.5	1516	4	US-09-549-106-3 Sequence 3, Appli
18	789.8	53.5	1516	4	US-09-550-394-3 Sequence 3, Appli
19	781	52.9	1505	4	US-09-193-377B-5 Sequence 5, Appli
20	780.2	52.8	1542	1	US-08-114-695A-1 Sequence 1, Appli
21	780	52.8	1542	2	US-08-757-653-158 Sequence 158, App
22	779.8	52.8	1474	1	US-08-114-695A-8 Sequence 8, Appli
23	778.4	52.7	1556	3	US-08-995-960-1 Sequence 53, Appl
24	772.6	52.7	1484	2	US-08-632-470-53 Sequence 53, Appl
25	770.6	52.2	1518	1	US-08-114-695A-6 Sequence 6, Appli
26	769.6	52.1	1502	4	US-09-187-946-2 Sequence 27, Appl
27	768.4	52.0	1485	1	US-08-299-810A-27

28	766.8	51.9	4403765	4	US-09-103-840A-2 Sequence 2, Appli
29	766.6	51.9	1542	2	US-08-875-445-21 Sequence 21, Appl
30	754.6	51.1	1494	2	US-08-632-470-49 Sequence 49, Appl
31	754.2	51.1	1452	2	US-08-642-229A-2 Sequence 2, Appli
32	753.6	51.0	1455	2	US-08-642-229A-3 Sequence 3, Appli
33	749.6	50.8	1464	2	US-08-938-858-1 Sequence 1, Appli
34	749	50.7	1508	2	US-08-632-470-44 Sequence 44, Appl
35	748.4	50.7	1428	4	US-09-193-377B-1 Sequence 1, Appli
36	747.8	50.6	1452	1	US-08-276-943-1 Sequence 1, Appli
37	747.8	50.6	1452	2	US-08-716-841-1 Sequence 1, Appli
38	745.8	50.5	1654	4	US-09-216-909-1 Sequence 1, Appli
39	745.8	50.5	1654	4	US-09-702-843-1 Sequence 1, Appli
40	745.8	50.5	1654	4	US-09-702-843-1 Sequence 1, Appli
41	744.6	50.4	1430	2	US-08-902-518A-1 Sequence 1, Appli
42	743.8	50.4	1415	2	US-08-632-470-52 Sequence 52, Appl
43	742.4	50.3	1503	4	US-08-943-571-1 Sequence 1, Appli
44	741.8	50.2	1440	2	US-08-632-470-26 Sequence 26, Appl
45	741.8	50.2	1441	4	US-09-193-377B-6 Sequence 6, Appli

#### ALIGNMENTS

RESULT 1  
US-08-642-229A-1  
Sequence 1, Application US/08642229A  
Patent No. 5874291  
GENERAL INFORMATION:  
APPLICANT: Herwig, Russell P.  
APPLICANT: Bielefeldt, Angela R.  
APPLICANT: Stensel, H. David  
APPLICANT: Strand, Stuart E.  
TITLE OF INVENTION: Degradation of Environmental Toxins by a  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Christensen O'Connor Johnson and Kindness PLLC  
STREET: 1420 Fifth Avenue, Suite 2800  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: WA 98101-2347  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/642,229A  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/246,865  
FILING DATE: 20-MAY-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Shelness, Diana K.  
REGISTRATION NUMBER: 35,356  
REFERENCE/DOCKET NUMBER: UOFW19233  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 682 8100  
TELEFAX: (206) 224 0779  
TELEX: 4938023  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1536 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: nucleic acid  
DESCRIPTION: "16S ribosomal DNA"  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:

ORGANISM: Comamonas testosteroni ATCC No. 5874291 11996  
US-08-642-229A-1

Query Match 55.1%; Score 814.2; DB 2; Length 1536;  
Best Local Similarity 74.7%; Pred. No. 5.2e-284;  
Matches 1113; Conservative 0; Mismatches 343; Indels 33; Gaps 6;

```

Qy 1 gatacgtgctgaagaactaacgctggcgccgtcttaacacatgcaagtcgaagcggtgag 60
Db 16 GATCCTGGCTGACATTAACAGCTGGCGCATGCTTTACACATGCAAGTCCAGACGCTTACA 75
Qy 61 caa-----tacctacgagcaacgggtgtagtaaccgctgtatctctccgag 110
Db 76 GGTCTTGATGCTGACGAGTGGCGAACGGGTGAGTATATACATCGGAACGTGCTTAGTAG 135
Qy 111 tctggaataccttcgcgaagaagaactaataccggaatgctctgttgcatacaagatt 170
Db 136 TGGGGGATTAACCTACTCTCAAGAGATAGCTAATACCGCATGACATCTACGGATGAAGCA- 193
Qy 171 tgaaggtaaagatttattgtcttggaatgtagcccgccgagcttaagctagttgtgag 230
Db 194 GGGGACCTTGGCGCTTGTGCTACTAAGAGCGGCTGATGGAGATGAGTATGATGTGGGG 253
Qy 231 taatggtcacaagaagcgagcatcgtagccgctgagaaaggtgctcgcgcacaaatgga 290
Db 254 TAAAGGCTTACCAAGCGCGCATCTGTAGTGTGAGAGAGACAGACGACACTGAGG 313
Qy 291 actgagacaggttcataactcctcaggaagcgagcaatgaatcttgcataatggggg 350
Db 314 ACTAGACACGCGCCAGCTCTACGGGAGCAGCAGTGGGAAATTTGGCAATGGGCG 373
Qy 351 aaaccgtgaagccgagccgctggaacgaagaagtcctcgatgtglaaagttcata 410
Db 374 AAAGCTGATCAGCAATGCGCGCTGCAAGATGAAGGCCCTCGGGTTTATAACGCTTTT 433
Qy 411 ggcagaaataaagca-----gcaatgtgataatgtagtacctgacct--aag 455
Db 434 GTACGGAAAGCAAAAGCGCGCTAATATCCCGGTCATACCGTACCGTAAGAAATGAG 493
Qy 456 caccggcctaactaagtgcgaacgacggcggaataatgtagtgcgaacgctgttcggaa 515
Db 494 CACGGCTTAACCTAGCTGTCACAGCGCGGTATTAACGTAAGGCTTTAATCGGAA 553
Qy 516 tcaatggcgctaaaggggtgtagcggaatctgtaagcaagtcgtaaacgtcgagctc 575
Db 554 TTACTGGCGCTAAAGCGTCCAGCGGCTTTGTGAAGACAGTGGTAATCCCGGCTC 613
Qy 576 aaccgtgacctgcaactgtaaacatacaagtcgtgagattgggagaagcgaaatggaatcc 635
Db 614 AACCTGGGAACCTGCCATTTGTGACTCAAGGCTAGAGTCCGCGACAGGGGATGGAATTC 673
Qy 636 aaggtgtaacggtgaatgctgtagatatctggaagaacacgctggcggaagcgacttgct 695
Db 674 GCGTGTACAGGAATCGTATGATATCCGGAGAACACCGATGCGGAAGGGAATCCCT 733
Qy 696 ggcctaaacatgacgctgtagcaacgaagcgctggtaglaaacsagattatgatacccg 755
Db 734 GGGCTGACATGACGCTCATGACAGAAAGCGTGGGAGCAAAACAGATTAGTATACCTGG 793
Qy 756 taatcacgcacctaaagctgtctaccagttgttgagggtttaaaccctcagtaacgaac 815
Db 794 TAGTCCAGGCCCTTAAACATGTCATCTGTTGTTGCTTAACTGACCTACATTAACGAG 853
Qy 816 ctaagaggttaagtagaacgcctggggaactatgctcgcaagaatgaactaaagaatt 875
Db 854 CTAACGCGTGAAGTGTGACCGCTGGGAGTACGGCCGACAGTTTAAACTCAAGAAATTT 913
Qy 876 gacgggggttcgcacaaagcggtgagacatggtttaaattcagatgatacccaaaacct 935
Db 914 GACGGGGACCGCCACAAAGGGTGAATGATGTTTAAATTCATGCAACGCAAAACCT 973
Qy 936 caactgagcttgacatgtagctgtaatactgtagagatatatgagccttcggcgaga---- 991
```

```

Db 974 TACCCACCTTTGACATGAGCGAGAACTTACCAGATGATGTTGTGCTCGAAGAGAACT 1033
Qy 992 -ttcaaggtgctgcatggtgttcgctcaagctgctgcgtgagatgtgttggttaagtcgg 1050
Db 1034 GCACACAGGTGTGCAATGAGCTGTGTCAGCTGTCGTGAGATGTTGGGTAAATCCCG 1093
Qy 1051 caacgaagcgaaacccctatcgatgtgttgcacctaaagtgtggcactgtagtaacgaactgc 1110
Db 1094 CAACGAGCGCAACCTTGCCATTAGTTGCTACATTCAGTTGACACTCTAATGGAGCTGC 1153
Qy 1111 cggltgacaaacgggaaggaagcggaatgacgtcaaatcccaatgagccttattgtccagg 1170
Db 1154 CGGAGACAAACCGGAGGAAGGTGGGATGACGTCAAGTCTCATAGGCCCTTAAAGTGGG 1213
Qy 1171 gccacaagctgctacaatgtagccgatacaagaggtgacgaactcgcaagaaggaactaat 1230
Db 1214 GCTACACAGCTCATACATGAGCTGTGTCAAAAGGTTGCCAAACCGGAGGGAGCTAAT 1273
Qy 1231 ctctaaagctcgtgccagctgcatgtgggtctcgaactcgaccccatgaagtgcgaat 1290
Db 1274 CCCATTAAGCCAGTCATGATGCTGATCCGATGCAAGTGTGCAACTGCGTGAAGTGGAA 1333
Qy 1291 cgtcagtaatcgcggaatcagacatgacggtgtagtaacgttcccggaactgtgacacgg 1350
Db 1334 CGCTAATATCTGTGATCAAAATGTCAGGGTGAATACGTTCCGGGTGTACACACG 1393
Qy 1351 cccgtcacacacactgagtggtggagccacggaaatggtctctgtgtaacgtaaggaaga 1410
Db 1394 CCCGTACACCATGAGGAGCGGCTCTCCGCAAGATAG-GRAGCCTAACCGTAAGGAGGCG 1452
Qy 1411 gactactaaggtgaactcgtaaaagggtgtagaactgtaacgaagttcc 1459
Db 1453 GCTTACCAACGCGGGGTGCTGTGACTGGGCTGTAAGTCTTACCAAGTAGC 1501
```

```

RESULT 2
US-09-198-955A-14
: Sequence 14, Application US/09198955A
: Patent No. 6187580
: GENERAL INFORMATION:
: APPLICANT: Andersen, Lene N.
: APPLICANT: Schlein, Martin
: APPLICANT: Lange, Niels E.
: APPLICANT: Bjornvad, Mads E.
: APPLICANT: Moller, Soren
: APPLICANT: Glad, Sanne O. S.
: APPLICANT: Kaupinen, Markus S.
: APPLICANT: Schmitt, Kirk
: APPLICANT: Kongsbak, Lars
: TITLE OF INVENTION: No. 6187580el Pectate Lyases
: FILE REFERENCE: 5378.200-US
: CURRENT APPLICATION NUMBER: US/09/198, 955A
: PRIOR FILING DATE: 1998-11-24
: PRIOR APPLICATION NUMBER: 1343/97
: PRIOR FILING DATE: 1997-11-24
: PRIOR APPLICATION NUMBER: 1344/97
: PRIOR FILING DATE: 1997-11-24
: PRIOR APPLICATION NUMBER: 60/067, 249
: PRIOR FILING DATE: 1997-12-02
: PRIOR APPLICATION NUMBER: 60/067, 240
: PRIOR FILING DATE: 1997-12-02
: PRIOR APPLICATION NUMBER: 09/073, 684
: PRIOR FILING DATE: 1998-05-06
: PRIOR APPLICATION NUMBER: 09/184, 217
: PRIOR FILING DATE: 1998-11-02
: NUMBER OF SEQ ID NOS: 32
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 14
: LENGTH: 1508
: TYPE: RNA
: ORGANISM: Bacillus sp.
US-09-198-955A-14
```



TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1521 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: *Bacillus* sp.  
INDIVIDUAL ISOLATE: DSM 8722D  
US-08-501-126-20

Query Match	54.5%	Score 805.4	DB 3	Length 1521
Best Local Similarity	76.9%	Pred. No. 7.9e-281		
Matches 1097	Conservative 0	Mismatches 292	Indels 37	Gaps 8

[illegible]

RESULT 4  
 / Sequence 7, Application US/08114695A  
 / Patent No. 5508193  
 / GENERAL INFORMATION:  
 / APPLICANT: Mandelbaum, Raphael T.  
 / APPLICANT: Mackett, Lawrence P.  
 / TITLE OF INVENTION: DEGRADATION OF S-TRIAZINES IN SOIL AND  
 / TITLE OF INVENTION: WATER  
 / NUMBER OF SEQUENCES: 8  
 / CORRESPONDENCE ADDRESSES:  
 / ADDRESSEE: SCHEGMAN, LUNDBERG & MOESSNER, P.A.  
 / STREET: 3500 IDS CENTER  
 / CITY: MINNEAPOLIS  
 / STATE: MN  
 / COUNTRY: USA  
 / ZIP: 55402  
 / COMPUTER READABLE FORM:  
 / MEDIUM TYPE: Floppy disk  
 / COMPUTER: IBM PC compatible  
 / OPERATING SYSTEM: PC-DOS/MS-DOS  
 / SOFTWARE: PatentIn Release #1.0, Version #1.25  
 / CURRENT APPLICATION DATA:  
 / APPLICATION NUMBER: US/08/114,695A  
 / FILING DATE: 31-AUG-1993  
 / CLASSIFICATION: 435  
 / ATTORNEY/AGENT INFORMATION:  
 / NAME: MUETING, ANN M.  
 / REGISTRATION NUMBER: 33,977  
 / REFERENCE/DOCKET NUMBER: 600.268051





CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/995,960  
FILING DATE: 22-DEC-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: RESNICK, David S  
REGISTRATION NUMBER: 34,235  
REFERENCE/DOCKET NUMBER: 1826/47986  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1512 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: rRNA  
US-08-995-960-2

Query Match 54.3%; Score 801.8; DB 3; Length 1512;  
Best Local Similarity 75.3%; Pred. No. 1.6e-279;  
Matches 1134; Conservative 0; Mismatches 322; Indels 49; Gaps 9;

QY 1 gataatgctcagaactaagctgagcgagcgtcttaaacatgcaagtcgagcgagtag 60  
DB 8 GATCATGCTCAGAGATGAACCTGCGCGCTGCTTAATGCAATGCAAGCGAGCAATGA 67  
QY 61 caat-----acctaagcgagcgagcgagtagtaaacgt-gtaatcttc 103  
DB 68 TTAAGAGCTTCTCTTATGAGTTAGCGCGAGCGGCTAGTAACCTGGTAACTAC 127  
QY 104 ctccgagctcgtgaatactcctcgaagaagtaataccgagtagctcgttgaatca 163  
DB 128 CCATTAAGACTGGATTAACCTCGGGGAACCGGGGCTAATACCGGATTAATTTGAATGC 187  
QY 164 caagattgataagtaag-----attatgcttgagagatgagccgcgagcagata 216  
DB 188 ATAGTTCCAAATTTGAAAGCGCGCTTGGCTGTCACTTATGATGAGGACCGCGCTCCATTA 247  
QY 217 gctagttgtagtgaatgagctcaccagaagcgagatcgtagcgagcctagaaggtgt 276  
DB 248 GCTAGTTGGTGGTAAAGGCTCACCAAGGCGACGATGCGTAGCCGACCTTGAAGGTGA 307  
QY 277 ccgagcacaatggaacttgagacacggtccatactcctcgaagagcgagcagttaaagaatc 336  
DB 308 TCGGCGACACTGGGAGCTAGACACGGCCAGACTCTCAAGGAGGACAGTAAAGGATTC 367  
QY 337 ttgctcgaatggggaagaccctggaagcagcgagcgagtgagcgaagaaggtctcgagat 396  
DB 368 TTCCCGCAATGGACGAAGCTGTACGAGCAACCGCGCTGATGATGAAAGGCTTCGGGT 427  
QY 397 tgaagttcatatgagcaggaataaataagcagcaatgtgatgatgtactcgtcct----- 451  
DB 428 CGTAAACTCTTCTTTAGGGAAGAAACAAGTCTAGTTGAATAACTGGCACCTTGAGCGG 487  
QY 452 -----aaagcaacgagcctaactacgtagcagcag-ccgaggtataacgttggtag 499  
DB 488 TACTTAACCAACCAACCAACGCTTAACGTCAGCCAGACCGCGCTAATACGTAGGTGG 547  
QY 500 caagcgtgttcggaatcatcttgagcgttaagaggtgtaggaggttctgtaagtagtga 559  
DB 548 CAAGCGTTATCCGGAATTAATGGCGTAAAGCGCGGACGGGTCTTAAAGTCTGATG 607  
QY 560 tgaatactgagcgcaacccgctgagcctgagcacttgaatacagaatctgaggttggag 619  
DB 608 TGAAGCCCAACGGCTACACCGTGGAGGCTGATTTGGAATCTGGGAAGACTTGAGTGGCAGAG 667

QY 620 aggcgaatggaattccaggtgtagcgttgaataatcgtaagatcttgtaggaacacagtg 679  
DB 668 AGGAAGTGGAAATTCATGATGTAGCGGTGAATGCGTAGAGATATGAGGAACACCAAGTG 727  
QY 680 ggcgaagcgagcttgcgtgctcgaactgaagcgtgaggaagcggtggtgtagtaag 739  
DB 728 GCCAAGCGCACTTCTGTGTACTGACACTGAGCGCGGAAGCGTGGGAGCAAAACA 787  
QY 740 ggaatgatacccccgaatccagccctaaacgttcttaccagttgttgggtttc- 798  
DB 788 GGATTAGATACCTCGTAGTGCACGCCGTAAAGATGAGTCAAGTGTAGAGGGTTTC 847  
QY 799 -aacccctgaatgaacgaactaagcagatgaagaccgcttggggactatgctcgcaaga 857  
DB 848 CGCCCTTATGCTGAAGTTAAGCATTAAGCATCTCCGCGGAGATTAAGCGCGCAAG 907  
QY 858 gtgaactcaaggaatgtagcgaggggtcgcgaagaagcggtgagcatgtgttaattcg 917  
DB 908 CTGAAGCTCAAAAGAAATTAAGCGGGGCCCGACAAAGCGGTGAGCATGTGTTAATTCG 967  
QY 918 atgatacccaaaaacctcaactcgtgagcttgacatgtagctgaatcagtagagatatag 977  
DB 968 AAGCAACGCCAAGAACCTTAACAGAGCTTGAACAT-CCTGTGAAAACTTAGAGATAGAGC 1026  
QY 978 agc-----ctcgggcaatcacaagtgctgcatggtgttcgtcagctgtgtcgtaga 1033  
DB 1027 TTYCTCTTGGGAGCAAGTACAGTACAGTGTGATGTTGCTGATGCTGCTGTGTGAGA 1086  
QY 1034 tgttggttaagtcgcgcgaagcagcaacccct-accgtatgttcttaacttaagtttg 1092  
DB 1087 TGTGGGTAAAGTCCCGCAAGACCGACACCTTGATCTTAAGTGCATTAATTAAGTTGG 1146  
QY 1093 gcaatgtagaagacgctgcggttgacaacacggaagagcgaggatgaatcattcaatcctc 1152  
DB 1147 GCACTTAAGGTGACTGCGGTGACAAACCGGAGAGGTGGGATACGTCGAATATC 1206  
QY 1153 atgaccttaatgctcagagccacacagctgttacaatgagccgagagaggtcccaac 1212  
DB 1207 ATGCCCTTATGACCTTGCGTACACAGTGTGCCAATGAGAGGTACAAAGCTGCAAGA 1266  
QY 1213 tcgcaagagagagtaactctcctaagaatcggtcccaatctcgatcttgaggttgcgaactcg 1272  
DB 1267 CCGCGAGGTGAGCTAATCTAATAAOCCTTCAAGTTCGAGTTGAAGCTGCAACTCG 1326  
QY 1273 acccatgaagtcggaatcgtagtaacgagatcagacatgcgaggttgataatcttc 1332  
DB 1327 CCFATCATGAAGCTGGAATCCCTAGTAATCGCGATACACATGACCGCGGTGAATACGTTCC 1386  
QY 1333 cggacctgttacaacacgcccgtcaacacacactgaagtgaggagaccggaagtgtctt 1392  
DB 1387 CGGGCTTGTACACACCGCCCGTCAACACGAGAGTTTGTAAACACCGGAAGTGGTGCG 1446  
QY 1393 gtaacgctlaagagacagactactaagtgtaaacctcgtaaaaggggtgtaagtcgtata 1452  
DB 1447 GTAACCTTATGAGACCGCAGCGCCTAAGGTGGAGCATGATGTGGGTGAAGTGTGATA 1506  
QY 1453 aggtta 1457  
DB 1507 AGGTA 1511

RESULT 6  
US-08-757-653-160  
; Sequence 160, Application US/08/757653  
; Patent No. 5843669  
; GENERAL INFORMATION:  
; APPLICANT: Kaiser, Michael W.  
; APPLICANT: Lyamichay, Victor I.  
; APPLICANT: Lyamichay, Natasha  
; TITLE OF INVENTION: Cleavage Of Nucleic Acid Using  
; TITLE OF INVENTION: Thermostable FEN-1 Endonucleases  
; NUMBER OF SEQUENCES: 190  
; CORRESPONDENCE ADDRESS:







```

; APPLICANT: Kim, Sang H
; TITLE OF INVENTION: Method for Producing Amide Compounds Using a Nitrite
; FILE REFERENCE: MSU 4.1-401
; CURRENT APPLICATION NUMBER: US/09/248,528C
; CURRENT FILING DATE: 1999-02-10
; EARLIER APPLICATION NUMBER: 60/083,485
; EARLIER FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1517
; TYPE: DNA
; ORGANISM: Bacillus sp.
; FEATURE:
; NAME/KEY: rRNA
; LOCATION: (1)..(1517)
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: Z26929
; DATABASE ENTRY DATE: 1998-07-02
US-09-248-528-2
```

```

Query Match      53.6%; Score 792.4; DB 3; Length 1517;
Best Local Similarity 75.0%; Pred. No. 4e-276;
Matches 1128; Conservative 0; Mismatches 326; Indels 50; Gaps 9;
```

```

QY 6 tggctgaagactaacgctgagcggtctttaaactgaagtcgagcggtgaagaata 65
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 tggctgaagactaacgctgagcggtctttaaactgaagtcgagcggtgaagaata 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 66 cc-----taagcgagacgggtgagtaacacgtgtatcttcccg-ag 110
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 gcttgcctcttaagtgtagcgagcggtgagtaacacgtgtatcttcccg-ag 120
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 111 tctgggaatacttccggaagaaagctaataccgagatgctctgttgatcaagaatt 170
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 121 actggaataacttcggaagaaacggaactaataccggaataacacggaatacgtgtt 180
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 171 tgataagtaag-----attatgtcttgagagtagagccgagcgatagtagt 222
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 181 tgggtgaagacgagctttagctgtacgtgacgagatggcgccgagcattagtagt 240
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 223 tgggtgaagacgagctttagctgtacgtgacgagatggcgccgagcattagtagt 282
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 241 tgggtgaagacgagctttagctgtacgtgacgagatggcgccgagcattagtagt 300
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 283 acaatggaactgagacaggttccatctcctacgagagcgagcaggttaagaatctgtctc 342
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 301 acaatggaactgagacaggttccatctcctacgagagcgagcaggttaagaatctgtctc 360
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 343 aatggggaagaaacccctgaagcaagcgagcggtgaaacgaaaggtcttcgagttttaa 402
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 361 aatggggaagaaacccctgaagcaagcgagcggtgaaacgaaaggtcttcgagttttaa 420
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 403 gttcatatagcagaaanaataagcagcaatgtga-----tgatgtactctg 448
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 421 gttcatatagcagaaanaataagcagcaatgtga-----tgatgtactctg 480
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 449 ccta--aagcagcgcttaactacgtgacgagcgagcggttaactgtgtgcaagcgt 506
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 481 acaagaaagaaacgagcttaactacgtgacgagcgagcggttaactgtgtgcaagcgt 540
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 507 tgttcggaatcatctggcgcttaagaggtgtcgtaagcgaggttgtaagtctgaagtcgtaaac 566
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 541 tgttcggaatcatctggcgcttaagaggtgtcgtaagcgaggttgtaagtctgaagtcgtaaac 600
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 567 tgcgagctcaaacccggtgagcgtgacttgaactgaactaagaagtcgtggaaggaagcaag 626
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 601 tgcgagctcaaacccggtgagcgtgacttgaactgaactgaagtcgtggaaggaaggaag 660
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 627 tggaaatccaggtgtatgagcgtgaatgtatagatcttgaggaagaaacacagtcgtggaag 686
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```

DB 661 cggaaatccacgtgtagcgttgaaatgcg tagaagatgttgaggaacacacagctgagcaag 720
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 687 cgacttgcgtgcctcaaaactgagcgtgaggaacgaggaagcgttgagtagtaaacggattag 746
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 721 cggctctcgtgcctgtaactgagcgtgaggaacgaggaagcgttgagtagtaaacggattag 780
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 747 ataaccggttaactgagcgtgcctcaaaactgagcgtgaggaacgaggaagcgttgagtagtaaac 804
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 781 ataaccggttaactgagcgtgcctcaaaactgagcgtgaggaacgaggaagcgttgagtagtaaac 840
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 805 cagtaagaaactaagcagtaagtagagcgtgcctgaggaactatgctgcgaagatgaaac 864
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 841 tagtgcgtgagcaaaacgagcgtgaggaactatgctgcgaagatgaaac 900
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 865 tcaaggaatgtgacggtgtccgcaaaacggttgaggaactatgctgcgaagatgaaac 924
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 901 tcaaggaatgtgacggtgtccgcaaaacggttgaggaactatgctgcgaagatgaaac 960
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 925 cccaaaaacctcaacctgagcttgaat-----gagatgaaactatgagagataatg 977
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 961 ggaagaaacttaacacagcttgaatccctccggaacacccctagagatagagcgttccct 1020
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 978 agccttcggtgagatcacaagctgctgcaatgtgtgtcgtcagctcgtgtgagatgtt 1037
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1021 ttcggtgaggtgagcaggtgagcaaggtgtgcaatgtgtgtcgtcagctcgtgtgagatgtt 1080
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1038 ggttgaatgcctgcaacgagcgtgagcaacccctatgctgtgtgctga-ccttaagtgtgagc 1096
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1081 ggttgaatgcctgcaacgagcgtgagcaacccctatgctgtgtgctga-ccttaagtgtgagc 1140
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1097 tggtaagaaactgcgtgtgacaaacggtgaggaagcgtggtgagtcgaatcgaatcgtg 1156
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1141 tctgaagtgactgcgtgcaaaagcgtgaggaagcgtggtgagtcgaatcgaatcgtg 1200
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1157 ccttattgcagaggtgcacacacagctgtcacaatggtcgtgacaggaaggtgtcccaactgc 1216
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1201 ccttattgcagaggtgcacacacagctgtcacaatggtcgtgacaggaaggtgtcccaactgc 1260
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1217 aagaggaactaactcctcaaaagctgcgtccagctgaggtgtggtgtcgaactgagcc 1276
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1261 gaggtgaggtgaggttcccaaaacacactcgtgaggtgtggtgtcgaactgagcc 1320
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1277 catgaagctggaatgcgtgtaactgcgtgagtcagcaatgcgtgagtcgaatgcgtg 1336
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1321 catgaagctggaatgcgtgtaactgcgtgagtcagcaatgcgtgagtcgaatgcgtg 1380
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1337 ccttgcacacacggtgcacacacacacacacacacacacacacacacacacacacacacac 1396
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1381 tcttgcacacacggtgcacacacacacacacacacacacacacacacacacacacacacac 1440
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1397 accgtta-aggaagacagactaagtgaaactcgtlaaaggggtgaaagtcgtgaacaag 1455
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1441 ccttgcaggtgaggtgaggtgaggtgaggtgaggtgaggtgaggtgaggtgaggtgaggtgag 1500
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1456 tacc 1459
   |||||
DB 1501 tagc 1504
```

```

RESULT 9
US-09-549-108-2
; Sequence 2, Application US/09549108
; Patent No. 6214603
; GENERAL INFORMATION:
; APPLICANT: Oriol, Patrick J
; APPLICANT: Padmakumar, Rugmini
; APPLICANT: Kim, Sang H
; TITLE OF INVENTION: Method for Producing Amide Compounds Using a Nitrite
; FILE REFERENCE: MSU 4.1-486
; CURRENT APPLICATION NUMBER: US/09/549,108
; EARLIER FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/083,485
```

PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 09/248,528  
PRIOR FILING DATE: 1999-02-10  
NUMBER OF SEQ ID NOS: 18  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 2  
LENGTH: 1517  
TYPE: DNA  
ORGANISM: Bacillus sp.  
FEATURE:  
NAME/KEY: rRNA  
LOCATION: (1)-(1517)  
DATABASE ACCESSION NUMBER: Z26929  
DATABASE ENTRY DATE: 1998-07-02  
US-09-549-108-2

Query Match 53.6%; Score 792.4; DB 4; Length 1517;  
Best Local Similarity 75.0%; Pred. No. 4e-276;  
Matches 1128; Conservative 0; Mismatches 326; Indels 50; Gaps 9;

QY 6 tggctcagaactaagcgtgagcggtcgttaacaatgcaagtcgagcggtgagcaata 65  
DB 1 tggctcagaactaagcgtgagcggtcgttaacaatgcaagtcgagcggtgagcaata 60  
QY 66 c-----tagcgagcgagcggtgagtaacacgtgtaactctcccg-ag 110  
DB 61 gctgctccttaagcttagcgagcggtgagtaacacgtgtaactctcccg-ag 120  
QY 111 tctggataacttcggaagaagataataacgagtagcgcgttgatcaacaagt 170  
DB 121 actggataacttcggaagaagataataacgagtagcgcgttgatcaacaagt 180  
QY 171 tgaatgtaag-----attatgtgtgagagtagagccgcgcgcgttagt 222  
DB 181 tggatgtaagcggtcttagctgtcactgagatggtgcccgcgcgcgttagt 240  
QY 223 tgggagagtagtggctcacaagcgagatcggttagcgcgcgcgcgcgcgcgcgc 282  
DB 241 tgggagagtagtggctcacaagcgagatcggttagcgcgcgcgcgcgcgcgcgc 300  
QY 283 acaatgagacgagacagcggtcactactcactcagagcgagcgagtaagaatctgctc 342  
DB 301 acaatgagacgagacagcggtcactactcactcagagcgagcgagtaagaatctgctc 360  
QY 343 aatgggggaaaccccgagcgagcggtgagcgagtaagaagtgcttgatgtaaa 402  
DB 361 aatgggggaaaccccgagcgagcggtgagcgagtaagaagtgcttgatgtaaa 420  
QY 403 gtcatatgagcaggaataaagcagcaatgtga-----tgatgtacctg 448  
DB 421 gctcgtgtgtcaggaagaagaagatgcttcgacaagcggtgacctgtgacctgt 480  
QY 449 ccta--aagcagcggtcactaactgagcgagcgagcggtgataatggtgcaagct 506  
DB 481 acgaagaagcgagcggtcactaactgagcgagcgagcggtgataatggtgcaagct 540  
QY 507 tgttgagatcgttgagcggtgagcggtgagcggtgagcggtgagcggtgagcggt 566  
DB 541 tgttgagatcgttgagcggtgagcggtgagcggtgagcggtgagcggtgagcggt 600  
QY 567 tggcgagcgaaccccggtgagcggtgagcggtgagcggtgagcggtgagcggtgag 626  
DB 601 tggcgagcgaaccccggtgagcggtgagcggtgagcggtgagcggtgagcggtgag 660  
QY 627 tggaaatccaggtgtgagcggtgagcggtgagcggtgagcggtgagcggtgagcggt 686  
DB 661 cgaatccaggtgtgagcggtgagcggtgagcggtgagcggtgagcggtgagcggtgag 720  
QY 687 cgaatccaggtgtgagcggtgagcggtgagcggtgagcggtgagcggtgagcggtgag 746  
DB 721 cgaatccaggtgtgagcggtgagcggtgagcggtgagcggtgagcggtgagcggtgag 780

QY 747 ataccggtatccagcgcccttaacgtgtctaccaggtgtgtggggtttt--aaccc 804  
DB 781 ataccggtatccagcgcccttaacgtgtctaccaggtgtgtggggtttt--aaccc 840  
QY 805 caatlaacgaaactaagtaagtaagtaagtaagtaagtaagtaagtaagtaagtaag 864  
DB 841 tagtgcgagcaaacgcatlaagcaatccgctggtgagtagcgcgcgcgcgcgcgcgc 900  
QY 865 tcaaggaatgagcggtgagcggtgagcggtgagcggtgagcggtgagcggtgagcggt 924  
DB 901 tcaaggaatgagcggtgagcggtgagcggtgagcggtgagcggtgagcggtgagcggt 960  
QY 925 ccaaaaacccctcctgagctgagct-----ggatctgaatcagtagagatag 977  
DB 961 gcaaggaacttaacggtgtctgacatccctcgaacacccctagagtagggtgtccct 1020  
QY 978 agcctcggcagatccagcggtgagcggtgagcggtgagcggtgagcggtgagcggt 1037  
DB 1021 ttcgggggagcaggtgagcggtgagcggtgagcggtgagcggtgagcggtgagcggt 1080  
QY 1038 ggggttagtccgcaagcgagcgagccctcctatgctgtgcta--cctaagttgggac 1096  
DB 1081 ggggttagtccgcaagcgagcgagccctcctatgctgtgcta--cctaagttgggac 1140  
QY 1097 tggatcgaactgagcggtgagcggtgagcggtgagcggtgagcggtgagcggtgagcggt 1156  
DB 1141 tctaagtgagctgagcggtgagcggtgagcggtgagcggtgagcggtgagcggtgagcggt 1200  
QY 1157 cctttagtccagcgagcgagcgagcggtgagcggtgagcggtgagcggtgagcggtgagcggt 1216  
DB 1201 cctttagtccagcgagcgagcgagcggtgagcggtgagcggtgagcggtgagcggtgagcggt 1260  
QY 1217 aagagggagcgaatcctcctaaagtcggtcccaagtcggtgagcggtgagcggtgagcggt 1276  
DB 1261 gagggtgagcgagcgaatcctcctaaagtcggtcccaagtcggtgagcggtgagcggtgagcggt 1320  
QY 1277 catgaagtcggaatcgtgagtaagtcggtgagcggtgagcggtgagcggtgagcggtgagcggt 1336  
DB 1321 catgaagtcggaatcgtgagtaagtcggtgagcggtgagcggtgagcggtgagcggtgagcggt 1380  
QY 1337 cctttagtccagcgagcgagcgagcggtgagcggtgagcggtgagcggtgagcggtgagcggt 1396  
DB 1381 tctgtgacaccccggtcagcagcgagcggtgagcggtgagcggtgagcggtgagcggtgagcggt 1440  
QY 1397 acgcta--agagagagcactaaggtgagcggtgagcggtgagcggtgagcggtgagcggt 1455  
DB 1441 cctttagtccagcgagcgagcgagcggtgagcggtgagcggtgagcggtgagcggtgagcggt 1500  
QY 1456 tacc 1459  
DB 1501 tagc 1504

RESULT 10  
US-09-549-111-2  
Sequence 2, Application US/09549111  
Patent No. 6228633  
GENERAL INFORMATION:  
APPLICANT: Oriol, Patrick J  
APPLICANT: Padmakumar, Rugmini  
APPLICANT: Kim, Sang H  
TITLE OF INVENTION: Method for Producing Amide Compounds Using a Nitrile  
FILE REFERENCE: MSU 4.1-489  
CURRENT APPLICATION NUMBER: US/09/549,111  
PRIOR FILING DATE: 2000-04-13  
PRIOR APPLICATION NUMBER: 60/083,485  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 09/248,528  
PRIOR FILING DATE: 1999-02-10  
NUMBER OF SEQ ID NOS: 18  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 2



DATABASE ACCESSION NUMBER: 226929  
 DATABASE ENTRY DATE: 1998-07-02  
 US-09-549-106-2

Query Match 53.6%; Score 792.4; DB 4; Length 1517;  
 Best Local Similarity 75.0%; Pred. No. 4e-276;  
 Matches 1128; Conservative 0; Mismatches 326; Indels 50; Gaps 9;

```

Qy 6 tggctcaagaactaacgttgcgcgcgtcttaacaatgcagtcgagcgagtagcaata 65
Db 1 tggctcaagaactaacgttgcgcgcgtcttaacaatgcagtcgagcgagtagcaata 60
Qy 66 c-----tagcgcgcaacggtgtagtaacacgttgtaactcttcctcg-ag 110
Db 61 gcttgctccttaggttaggcgcgcgaggtgtagtaacacgttgtagcaactgcctcgag 120
Qy 111 tctgggaataacttcggaagaagagtaataacgagtagtctctgttgatcaacaagatt 170
Db 121 actgggaataacttcggaagaagagtaataacgagtagtctctgttgatcaacaagatt 180
Qy 171 tgaatagtaag-----attatgtcttgagatgagcccgcgagcagattagctagt 222
Db 181 tctggttgaagaagcgctttagctgttcaactgcagatgagcccgcgagcagattagctagt 240
Qy 223 tggtagagtaatggtctcaacaagcgagatcgtagccgctcgagaggtgtccggcc 282
Db 241 tggtagagtaatggtctcaacaagcgagatcgtagccgctcgagaggtgtccggcc 300
Qy 283 acacttggaactgcagacacggtctactctcaacgagcgagcagattgaagatctgtc 342
Db 301 acacttggaactgcagacacggtctactctcaacgagcgagcagattgaagatctgtc 360
Qy 343 aatgggggaaacccctgaaacgagcgagcgtagaagaagaaggtcttgatgtgtaa 402
Db 361 aatggagaaagctgcagcgagcgagcgtagaagaagaaggtcttgatgtgtaa 420
Qy 403 gttcaatgagcagaagaataatgacgaattgta-----tgaatgtacctg 448
Db 421 gttctgtgtcagaggaagaacaagtaacgttgcgaacgagcggtatccttgacgttacctg 480
Qy 449 ccta--aagcacccgactactactgcagcagcgcgtaataacgtatgtagcaagcgt 506
Db 461 acgaggaagacacggtctactactgcagcagcgcgtagaatacgtatgtagcaagcgt 540
Qy 507 tgttcggaatacatgtggcgtagaaggtgtcgtagcgagatttgtagtaagttgtaaac 566
Db 541 tgttcggaatacatgtggcgtagaaggtgtcgtagcgagatttgtagtaagttgtaaac 600
Qy 567 tgcgggctcaaccgctggcctgcacttgaaactacaaagttctggagagagcaag 626
Db 601 tgcgggctcaaccgctggcctgcacttgaaactacaaagttctggagagagcaag 660
Qy 627 tgggaatlcacaggtttagcggtgaaatgcgtatgatactcggaagaacacaggtggcgag 686
Db 661 cggaaatlcacaggtttagcggtgaaatgcgtatgatactcggaagaacacaggtggcgag 720
Qy 687 cgaacttgcggtctcaaaactgcagtcgagcaggaagcgtaggtgtagtaaaagagattag 746
Db 721 cggcctctcgcgctgtaactgcagtcgagcaggaagcgtaggtgtagtaaaagagattag 780
Qy 747 ataccgcgtaatlcacagccttaaacgttgcacaaagttgttgagggtttt--aacct 804
Db 781 ataccgtgttagtccagcgcgttaaacagtagttagttagtgtaggggtatccacctt 840
Qy 805 cagtaacgaactaacaggttagtagaacgcgtcgtaggagctatgtcgcgaagaagtagaac 864
Db 841 tagtgcctgagcaaaacgacgtctgcgctggggtgtagcagcgagcgagcgtagaac 900
Qy 865 tcaaaagaaattgacggggtccgcacaaagcggtgtagcagctgtgatttaattcgaatgatac 924
Db 901 tcaaaagaaattgacggggtccgcacaaagcggtgtagcagctgtgatttaattcgaatgatac 960

```

```

Qy 925 cccaaaacacctaacctgggcttgacat-----gactgtaatcatgtagaatatacg 977
Db 961 gcgaagaaccttaaccggtcttgacatccctcgacaacacctgaaataggggttccct 1020
Qy 978 agccttcgggcagatcatcaaggctgctgacatggtgtcgtcagctcgtgtcgtagatgtt 1037
Db 1021 ttcggggggacaggggtgacaggtgtgtgacatggtgtcgtcagctcgtgtcgtagatgtt 1080
Qy 1038 gggtagtcccgcaacgagcgcaacaccttaacgtatgtgtgta--ccttaagtgggacac 1096
Db 1081 gggtagtcccgcaacgagcgcaacaccttaacgtatgtgtgta--ccttaagtgggacac 1140
Qy 1097 tggtagaacttcggtgtagcaaacggaggaagcggggtagacgtcaaatcctatg 1156
Db 1141 tctaaagttagctgcggtctcaaaagtctgaggaaggtggggatgtagctcaaatcatgctc 1200
Qy 1157 ccttatgtccaaaggccacaacacgtgtctcaaatgtgcgagatacagaggttcgcgaactgc 1216
Db 1201 ccttatgtacgttggtctacacacgtgtctcaaatgtgtgtgtagcaaaaggcgagcgaac 1260
Qy 1217 aagagggagcttaactctcaaaagtctgctccagctcgagattgggtctgcaactcgaccc 1276
Db 1261 gaggtagaggaatcccaaaacactctcagttcgatgtgagagctgtgaactccgctg 1320
Qy 1277 catgaatctggaatcgtagtaatcgcggaatcagcatgcccgcggtgaaatcgttcccgga 1336
Db 1321 catgaagcggaaatcgtagtaatcgcggaatcagcatgcccgcggtgaaatcgttcccgga 1380
Qy 1337 cctgttaacacacgcccgtctacacacactgtgagtgaggagacccgaggtgtcttgta 1396
Db 1381 tctgttaacacacgcccgtctacacacactgtgagtgaggagacccgaggtgtcttgta 1440
Qy 1397 acgcta-aggagacagactactaaggtgaacactgtgaaggggtggaagtcgtgaacaagg 1455
Db 1441 ccttaagggagagcagcgccgcggaaggtgaggacaatgattgggtgtagcgtgaacaagg 1500
Qy 1456 taacc 1459
Db 1501 tagc 1504

```

```

RESULT 12
US-09-550-394-2
; Sequence 2, Application US/09550394
; Patent No. 6287828
; GENERAL INFORMATION:
; APPLICANT: Oriol, Patrick J
; APPLICANT: Padmakumar, Rugmini
; APPLICANT: Kim, Sang H
; TITLE OF INVENTION: Method for Producing Amide Compounds Using a Nitrite
; FILE REFERENCE: MSU 4.1-488
; CURRENT APPLICATION NUMBER: US/09/550.394
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/083.485
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 09/248.528
; PRIOR FILING DATE: 1999-02-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1517
; TYPE: DNA
; ORGANISM: Bacillus sp.
; FEATURE:
; NAME/KEY: IRNA
; LOCATION: (1)..(1517)
; DATABASE ACCESSION NUMBER: 226929
; DATABASE ENTRY DATE: 1998-07-02
; US-09-550-394-2

```

Query Match 53.6%; Score 792.4; DB 4; Length 1517;









Oy	286	atggaatctggaacacggtccatactctctcgggagcgacgataagaattcttgcataat	345
Db	301	ctggactctgaaacacggtccacagatctctccgggaagcgacgaattaggaattcttcgcataat	360
Oy	346	gggggaacacctgaagacgacgcgcgtctgaaacgaagaaggtctctcgattgtataaagt	405
Db	361	ggacgaagaagctctgacggaggaacacgcgcgttgagcgagaagaaggtctctcgatcgttaagct	420
Oy	406	catlaagcgaggaataaataagcagaatgtga-----tgatgtaacctgtcct	451
Db	421	ctgtgtctcaaggagaagaacaaagtctcgctctggaacacgggcgttaacctgtgcggtlactgtacg	480
Oy	452	a--aagacacgggtctaaactaagctgctgacgcacgcgcgttaatacgtatagttgcaagctgtt	509
Db	481	aggaagaacgaagcttaactaactaagctgctcgcgcgcgcgcgttaatacgtatggtgtgcgaacgctgtt	540
Oy	510	tcggaaatcatctggcgtlaaagaagtgctgcgtagcggaatttgtaagtcagaagtggtgaataactgc	569
Db	541	ccggaaatcatctggcgtlaaagaacgcgcgcgcgcgcgtctctcttaagctctgattgtgaataactgc	600
Oy	570	gggtctcaacccggtgacctgtgcacttgaaactataaagctctggaattctgggaagagcgaaagtgtg	629
Db	601	cggtctcaacccggtgacctgtgcacttgaaacttggaacttggaattctgggaagagcgaaagtgtg	660
Oy	630	aattccacggtgtctagcggctggaatactgtaagatactctggaagaacacacaaatctggcgaagaagcga	689
Db	661	aattccacggtgtctagcggctggaatactgtaagatactctggaagaacacacaaatctggcgaagaagcga	720
Oy	690	cttgctgtgctccaaactctgaacgctctgaggtcacgcaagaagcgtgtgattagtaaacgggattagata	749
Db	721	ctctctgtgctccaaactctgaacgctctgaggtcacgcaagaagcgtgtgattagtaaacgggattagata	780
Oy	750	ccccgggttaactcaacgccccttaaacgctgtgtctacaacagctgtgtctggggcttt--aacctcag	807
Db	781	ccccgggttaactcaacgccccttaaacgctgtgtctacaacagctgtgtctggggcttt--aacctcag	840
Oy	808	taacggaaccttaacggaattaaagttaagtaacgcgcgcctggggagctatactgctcgcgaagaagtgaataactca	867
Db	841	tgctgcgcgcgaacacgcatctaaagcatctccgcctcgcctggcggtgagctacgcgcgaagaagtgaataactca	900
Oy	868	aaggaatattggcgggtggtctccgcgaagaacggttgtagacatgtgtgtattatctcgatagatacccc	927
Db	901	aaggaatattggcgggtggtctccgcgaagaacggttgtagacatgtgtgtattatctcgatagatacccc	960
Oy	928	aaaaaacctccacttggtgctctgacatactggaactcgaatc---atgtagaatatatagaacct	982
Db	961	aaggaactcttaacggtctctgacatactccactcgcgaacacccctagagataaggtggcttccctctc	1020
Oy	983	tcgggacgaattcaacaggtctctgcaatgtgtctgctgcgtcaagctcgtgtcgtgtagatagtttgggtt	1042
Db	1021	gggggaacacgggttgtagaagaagttgtgtcgaatgtgtctgctgcgcgtcgtgtcgtgtagagatgttgggtt	1080
Oy	1043	aagctccgcgcgaacgacgcgaacacccctctgctgatactgtccta--ccttaagttttgggcactgtgta	1101
Db	1081	aagctccgcgcgaacgacgcgcgaacacccctctgacttattgtgcagcatctcagtttgggcactctaa	1140
Oy	1102	cgaaaactgtccggtgtgacaaacccggaggaagcggtggatgacgttcaactccatctgagctctt	1161
Db	1141	ggttacttgcgcgtctaaagaagctcggaggaaggttgggggtatgacgttcaaatcatcatctgcccctt	1200
Oy	1162	atgtccacgggcgaacacacgctgctatacatgtgcgaatacgaagaaggtctgcgaacactcgcgaagaag	1221
Db	1201	atgtccacgggcgtatacacacgctgctatacatgtgtgtgtacaaagggcgaagcgaaacacgcggaagt	1260
Oy	1222	ggagactaactctaaagaagctgcgtcccaagtctcggaattgtgggtctggaactcgaaccccatga	1281
Db	1261	ggagcgaaatcccaaaaaaacacactctcagttctggaattgtcgaagctcgcgaactcgcctcatgta	1320
Oy	1282	agtcggagatcgctatgattatctgctggatcagatcagatgacgcggttgataatcgttcccggaaccttg	1341
Db	1321	agccggagatcgctatgattatctgctggatcagatcagatcgcggttgataatcgttcccggaaccttg	1380

```

QY      1342 taccacccgcgttaaccacttagtgggagaccccgaagtgtcttgttaacgt 1401
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      1381 taacaacgcccgctcacaccacagagatttgtaaacccccgaagtcggtygglaacct 1440

QY      1402 a-aggagaagactctaagtgtaaactcgttaaagggggtgaagtcgtatacaaggatcc 1459
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      1441 acggagaccgacgcccgcaaagtgtyggacaatagatgtggggtgaagtcgtatacaaggatgc 1499


RESULT   15
US-09-549-108-3
; Sequence 3, Application US/09549108
; Patent No. 6214603
; GENERAL INFORMATION:
; APPLICANT: Ortel, Patrick J
; APPLICANT: Padmakumar, Rugmini
; APPLICANT: Kim, Sang H
; TITLE OF INVENTION: Method for Producing Amide Compounds Using a Nitrite
; FILE REFERENCE: MSU 4.1-486
; CURRENT APPLICATION NUMBER: US/09/549,108
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/083,485
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 09/248,528
; PRIOR FILING DATE: 1999-02-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1516
; TYPE: DNA
; ORGANISM: Bacillus pallidus
; FEATURE:
; NAME/KEY: rRNA
; LOCATION: (1)..(1516)
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: Z26930/genbank
; DATABASE ENTRY DATE: 1997-05-14
US-09-549-108-3


Query Match          53.5%; Score 789.8; DB 4; Length 1516;
Best Local Similarity 75.0%; Pred. No. 3,Se=275;
Matches 1124; Conservative 0; Mismatches 327; Indels 48; Gaps
9

```

Search completed: December 15, 2001, 03:22:44  
Job time: 7559 sec

```
QY 406 cattagcgagaaataagcagcaatgtga-----tgaatgtaacctgcct 451
Db 421 ctgtgtcagggagaaacagtgccgttcgaaacagggcggtacacctgacgtacacgtacg 480
QY 452 a--aagcaccgctaaactagctgcagcagccgcgtgaatacgtatgtgcaagcgttgt 509
Db 481 aggaagccacagcgtcaatactatgtgcagcagccgcgtgaatacgtatgtgcaagcgttgt 540
QY 510 tcggaatcattggcgtaaaaggtgcgtgaagcggtatgttaagtcaggctgtgaaaactgc 569
Db 541 ccggaattatgtggctaaagcgcgcagcggtctccttaagctgtatgtgaatctcg 600
QY 570 gggctcaaccgcgtgcctgcgtgaactgaacacaaagtcgtgaggtctgggagaggaatg 629
Db 601 cggctcaaccgcgcagccgtcctgtgaaacgtggggaactgtgagtcagagagggagcg 660
QY 630 aattccaggtgtcagcggtgaatltgctgaatalctgagagaaacacacagtcggcgaagcg 689
Db 661 aattccacgtgtcagcggtgaatltgctgaatalctgagagaaacacacagtcggcgaagcg 720
QY 690 ctgtcgtgcctaaactgaacgtgcagcagaaagcgttggtatgaacggtatagata 749
Db 721 ctctctgtgcctgaactgaacgtgcagcagaaagcgttggtatgaacggtatagata 780
QY 750 ccccggtaatccacgcgcctaaacgttgtctacagttgtgtggtgtt--aacccacg 807
Db 781 cccgtgtatgtcagcgcgtgaacgtatgagtggtatgaggtatccacccttag 840
QY 808 taacgaaacctaaacggtatgaatagtagacgcctggtgactatgctcgcaagatgaactca 867
Db 841 tgcgtgcagcaaacgcatlaagcacctgcgcctggtgagtagcgcgcaaggtctgaaaactca 900
QY 868 aaggaattgagcgggtgtccgcaacagcgttgagcatgtgtttaatcgtatgaaccc 927
Db 901 aaggaattgagcgggtgtccgcaacagcgttgagcatgtgtttaatcgtatgaacccg 960
QY 928 aaaaaacctcaacctgtgctgaacatgtgaatc-----atgtagagatatatgagcct 982
Db 961 aagaaaccttaaccaggtctctgtgacatccctctgacaacacctagagatagcgcttcccttc 1020
QY 983 tcgggcagatctcaacaggtgtcgtcatgtgttcgtcagctcgtgtcgtatggtgtgt 1042
Db 1021 ggggacaaggtgcagcaggtgtgtcagctgtgtcagctcgtgtcgtatggtgtgt 1080
QY 1043 aagtcgcgcaacgagcgcgaacccctatcgtatgtgtcga-ccttaagctggtgactgtga 1101
Db 1081 aagtcgcgcaacgagcgcgaacccctatcgtatggtgtcgaacatcagttggtgactgtga 1140
QY 1102 cgaaaactgcgcgtgtgaacaaacggaggaagcggtgaatgaactcaaaatccatggtcctt 1161
Db 1141 ggtgactgtgcgttaaaagtcggaaggtgtggtatgacgtcaaatcatcatgcccctt 1200
QY 1162 atgttcagggcgcacacacgtgtcgaatgcagatgcagaggtgtcgcgaactgcgaagag 1221
Db 1201 atgacacgtggcctacacacagctgtcgaatgcagaggtgtgtaaaagggcagcgaaggt 1260
QY 1222 ggaactaatctctaaagtcggttccacgttcggaattggtgtcgtgaactgcgaacccatga 1281
Db 1261 ggaagcaatcccaaaaaaaccaactcagttcgtatgtgcaggtgcgaactgcgtcatga 1320
QY 1282 agtcggaatcgtatgaatgcgtgaatgcagcatgcgcgtgtgaatacgttcccggaactgt 1341
Db 1321 agccggaatcgtatgaatgcgtgaatgcagcatgcgcgtgtgaatacgttcccggtctgt 1380
QY 1342 tacacacgcgcgttcacacacacgtatgtggtgagaccccggaagtggtcttgttaacgt 1401
Db 1381 tacacacgcgcgttcacacacacgtatgtggtgagaccccggaagtcggtgtggttaacccct 1440
QY 1402 a-aggagacagactactatggtgtgaactcgttaaaaggggtgaagtgcgttaacaaggtacc 1459
Db 1441 acgggagcgcgcgcgcggaaggtatggacaataatgtatggtgtgaggtcgttaacaaggtacc 1499
```



**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2001, 01:10:40 : Search time 4441.54 Seconds  
(without alignments)  
3573.426 Million cell updates/sec

Title: US-09-380-826A-1  
Perfect score: 1477  
Sequence: 1 gatcagctcgaactaac.....ccgtaacgattcctgcag 1477

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues  
Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: em\_estfun:\*  
2: em\_esthm:\*  
3: em\_estln:\*  
4: em\_estom:\*  
5: em\_estpl:\*  
6: em\_estba:\*  
7: em\_estro:\*  
8: em\_estov:\*  
9: em\_hlc:\*  
10: gb\_estl:\*  
11: gb\_estc:\*  
12: gb\_hlc:\*  
13: gb\_gss:\*  
14: em\_gss\_fun:\*  
15: em\_gss\_hum:\*  
16: em\_gss\_inv:\*  
17: em\_gss\_pln:\*  
18: em\_gss\_pro:\*  
19: em\_gss\_rtd:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Query Length	DB ID	Description
C 1	436.8	29.6	800	13	AQ960400 LERPD51TR
C 2	424.8	28.8	789	13	BH145112 TDGEP51TR
C 3	424	28.7	716	13	AQ957363 LERAP36TR
C 4	411.4	27.9	868	11	BF620647 HYSMC002
C 5	411	27.8	730	13	BH144909 TDGEQ40TR
C 6	407.4	27.6	720	13	CNS00WCV
C 7	392.8	26.6	755	13	BH016730
C 8	389	26.3	746	13	AQ957362 LERAP36TR
C 9	383.6	26.0	741	13	AQ957181 LERAP33TR
C 10	378.8	25.6	757	13	AQ957182 LERAP34TR
C 11	376.6	25.5	676	13	B73824 T274TR TAM
C 12	364.8	24.7	655	13	BH016557 TDGCN35TR

C 13	360.4	24.4	693	13	AQ956298	AQ956298 LERAJ19TR
C 14	359.6	24.3	643	13	BH144673	BH144673 TDGEN65TR
C 15	359.6	24.3	676	13	AQ969911	AQ969911 LERJ28TR
C 16	359.4	24.3	643	13	BH007227	BH007227 ee64e07.x
C 17	356.6	24.1	659	13	AQ969996	AQ969996 LERUR90TR
C 18	354.4	24.0	716	13	CNS01008	AL153840 Anopheles
C 19	352.4	23.9	677	10	AM221851	AM221851 EST298662
C 20	351.4	23.8	624	13	BH145348	BH145348 TDGEV44TR
C 21	350.6	23.7	616	13	AG019017	AG019017 Homo sapi
C 22	346.8	23.5	675	13	AQ956297	AQ956297 LERAJ19TR
C 23	343	23.2	603	10	AV531914	AV531914 AV531914
C 24	342.2	23.2	655	11	BI207639	BI207639 EST525679
C 25	340.4	23.0	768	13	AQ957272	AQ957272 LERAD54TR
C 26	340	23.0	637	13	AQ962145	AQ962145 LERGD59TR
C 27	339.6	23.0	687	13	AQ955848	AQ955848 LERAG49TR
C 28	339	23.0	632	13	AQ967817	AQ967817 LERIV15TR
C 29	337.4	22.8	618	13	AQ969912	AQ969912 LERUR28TR
C 30	336.6	22.8	601	13	B27166	B27166 T2C12TR TAM
C 31	335.6	22.7	775	11	BG645215	BG645215 EST506834
C 32	334.2	22.6	660	10	AM600903	AM600903 3B5 CDNA
C 33	333.6	22.6	592	13	AG019018	AG019018 Homo sapi
C 34	332.4	22.5	599	13	AQ968703	AQ968703 LERJ531TR
C 35	332	22.5	648	13	AQ957342	AQ957342 LERAP25TR
C 36	331.6	22.5	561	13	AG019935	AG019935 Homo sapi
C 37	331.6	22.5	616	10	BE356988	BE356988 DGI_146_B
C 38	331.6	22.5	673	13	AF075762	AF075762 AF075762
C 39	330.8	22.4	604	10	BE204135	BE204135 EST396811
C 40	327.6	22.2	662	10	AW736167	AW736167 EST332153
C 41	327.2	22.2	639	13	AQ956183	AQ956183 LERAI55TR
C 42	325.6	22.0	559	11	BF845685	BF845685 780B5 CDN
C 43	324.4	22.0	610	13	AQ961463	AQ961463 LERFL40TR
C 44	323.8	21.9	739	13	CNS01J02	AL146887 Anopheles
C 45	319.8	21.7	553	5	BE362383	BE362383 DGI_86_HO

#### ALIGNMENTS

RESULT 1  
AQ960400/c 800 bp DNA GSS 28-JAN-2000  
LOCUS LERPD51TR LERA Arabidopsis thaliana genomic clone LERPD51, DNA  
DEFINITION  
sequence.  
ACCESSION  
AQ960400  
VERSION  
AQ960400.1 GI:6788101  
KEYWORDS  
GSS.  
SOURCE  
Thale cress.  
ORGANISM  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidops.  
1 (bases 1 to 800)  
Buell,C.R., Lin,X., Pai,G., Barnstead,M., Bowman,C., Ullrich,T.,  
Feldblum,T., Liang,F., Creasy,T. and Fraser,C.M.  
Genomic survey sequencing of Landsberg erecta ecotype of  
Arabidopsis thaliana and identification of sequence-based  
polymorphisms  
Unpublished (2000)  
Contact: Xiaoying Lin  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: atetlgr.org  
For additional information, see <http://www.tigr.org/tdb/at/at.html>  
Similar to A. thaliana chloroplast sequence (GB:AP000423)  
Seq primer: TR  
Class: shotgun.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

#### FEATURES

source  
1. location/Qualifiers  
800  
/organism="Arabidopsis thaliana"  
/strain="Landsberg erecta"  
/db\_xref="taxon:3702"

```

/clone="LERFD51"
/clone.lib="LERA"
/note="Organ: Leaf; Vector: PHOS1; Total genomic DNA was
sheared to 0.9-1 kbp before ligation."
147 a 257 c 181 g 215 t
BASE COUNT
ORIGIN

```

Query Match	29.6%;	Score 436.8;	DB 13;	Length 800;
Best Local Similarity	74.9%;	Pred. No. 1.3e-109;		
Matches 589;	Conservative 0;	Mismatches 187;	Indels 10;	Gaps 3

OY	190	gcttggagatagaccgcgcgcgcgtctagctatctgtgtggaagtaatgctcaacgaagcgca	249
Db	780	gccccagagagggcgtccgcgtctgatttagcttggatgagagcaattaccttaccgaagcgca	721
OY	250	cgatcgtatagccgcgcgcgttgaagagggtgtccgcgcacaaltgaacttgagacacggtccat	309
Db	720	tgatctagatagctgtgtccgagagagatgatcaccctccactgtggacgtgaaacacggcccgac	661
OY	310	tccttaaggggaagcgcgcagcttaagaatctgtgtcaatgtggggaaaccccttgagcagcgag	369
Db	660	tccttaaggggaagcgcgcagcttggggaaattttccgcgaattggcgaaagccctgacgaagc	601
OY	370	ccgcgttgaacgaagaaggtctctcgatctgttaagltcaattatgagcaggaataataagcagc	429
Db	600	ccgcgttgaacgaagaagcctacgggtcctgaactcttctttccacag--agaaagaacaa	543
OY	430	aattgtatgatagtgtaccctgccttaagaaacgcgcgccttaactgaagtgccaacgcgcgta	489
Db	542	gacggatattctggggg-----atpaaagcgtcggcttaactcttctgccacagaccgcggtat	489
OY	490	acgtactgttgcgaacgcttgttgcgaatcaatcgtgggcgttaaaaggtgtcgtlaagcggat	549
Db	488	acagagagatgcacagcgttatccggaaatgatttggggcgttaaacgctctgtagtgactttt	429
OY	550	aagtcagttgtgaaaaactcgcgcctcaaccgcgtgtgcctgcgaacttgaactcaactctgtg	609
Db	428	aagtcggcgcgtcaaatccagagcgtcaaccccttgacacggcgtgaaactgacaaagcttg	369
OY	610	agtttggagagagcgaagtcgaatctccaggtgtlaagcgtgaagtgcgtatatacttgag	669
Db	368	agttgagagagagcgaagaggaattttccggtggagcgggtgaatggcgtagagatcggaag	309
OY	670	aaccaccagtgcgaaagcgcacttctgtgctcaaaaactgaagcgttgagcagcaagaagcgtgtg	729
Db	308	aacaccacacggcggaagacactctcgtggccgcacactgacactgagaaacgaagactag	249
OY	730	gttagtaaacgggatttagatacccggtlaatccacgcctcaaacgcttgcctcaacgattgt	789
Db	248	ggagcgaatagggaatttagatnaccacagtagtctctnaccgtaaaacgatggatctagggcct	189
OY	790	gggggttctaacc--tcagtaacgaaccttaacgaatgaatgaacgcgccttggggacctat	847
Db	188	gggcgctatgacacccgcgcgtcgtctagcttaacggcttaagatattcccccctgggagatgac	129
OY	848	ggttcgaagaagtgaaactctcaagaagtgtaacgggggttcgcgcacaagcgttgaagcattgtg	907
Db	128	ggttcgaagaagtgaaactctcaagaagtgtaacgggggttcgcgcacaagcgttgaagcattgtg	69
OY	908	gtttaaactgatgatcccccaaaactcaactacccttggtcgttgatctgatatgatcatgtta	967
Db	68	gtttatttgatgacaaaggaagaaaccttaccagggccttgacatgcgggaattcctctttaa	9
OY	968	gagata 973	
Db	8	aagaca 3	
RESULT	2		
LOCUS	BHL45112/c	789 bp	DNA
DEFINITION	TDG6S717H CTG Lycopersicon esculentum genomic clone CTG631121, DN		

ACCESSION	BHL45112
VERSION	BHL45112.1
KEYWORDS	GI:15200731
SOURCE	GSS.
ORGANISM	tomato.
	<i>Lycopersicon esculentum</i>
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta
	Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots;
	Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
	<i>Lycopersicon</i> .
REFERENCE	1 (bases 1 to 789)
AUTHORS	van der Hoeven, R., Sun, H., Cho, J., Utterback, T., Ronning, C. and
Tanksley, S.	
TITLE	Tomato Demethylated Genomic DNA Sequences
JOURNAL	Unpublished (2001)
COMMENT	Contact: CUCI
	University Genome Institute

Clemson University  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>  
tomato demethylated genomic DNA  
Insert length: 1270 Std Error: 0.00  
Seq primer: M13F-R  
Class: Shobgun: 10000011408

FEATURES	Location/Qualifiers
source	1. . 789

```

/organism="Lycopersicon esculentum"
/guiltivar="E6203"
/db_xref="taxon:4081"
/clone="CTG63IL21"
/clone_1b="cT0G"
/disease_type="young leaves"
/dev_stage="12-14 weeks post harvest"
/lab_host="E. coli JM109"
/notes="vector: p Bluescript SK(+); Site_1: EcoRI; Site_2:
XhoI; This library was made from short EcoRI digested
fragments of the genome of Lycopersicon esculentum ligated
into pBS (SK-). The fragments were cloned into the
methylation restrictive E. coli strain JM109 with the
purpose of enriching the library for non-methylated DNA
fragments. This procedure may enrich the pool of cloned
fragments in JM109 cells for sequences representing
expressed genes. Average insert size 1.27 kb."

```

Query Match	28.8%	Score	424.8	DB	13	Length	789
Best Local Similarity	76.2%	Pred No.	2.6e-106				
Matches	388	Conservative	0	Mismatches	177	Indels	7
						Gaps	5

Qy	546	ttgttaagtcacagttggaanaactcgcgggctctaaaccgcggtctgaacttgaactgaacaaagt	605
Db	782	TTTTTAAGTCGCCGCGCAATCCACGGGCTC-ACCTCGGACAGGGCGGTGAACACTACCAAG	724
Qy	606	ctcgagattctggagaagcgcaagtgaatctccagtgctagtcggtgaatctgcgaatctc	665
Db	723	CTGGAGTACGGTAGGGGCGAGAGAAATTTCCGGTGGAGACGGGTGAATCCGTAGAGATCGG	664
Qy	666	gaggaacaccagttgcgaagcgcgaactctcctgcctcaaacctgaacgcctggagccagaagaac	725
Db	663	AAGAACAACCAACGGCGAAGAGCACTCTCGTGGCGACACCTGAGAGAGAAAGC	604
Qy	726	gtgggttagttaaaccggatctaga taaccggcgaatccacgcgcctaaacgltctcaacagt	785
Db	603	TACGGGAGCGCAATTGGCATTAGTACCCTCCAGTAGTCTTAGCGCTTAACGATGAGATCAAGG	544
Qy	786	tgcttgggggttttaacc--ctagtaacgaacctaaaggaatgaagtagaaccgccttggga	843
Db	543	CGCTTGCGCTATCGACCCCTGCAGTCTGTAGCTAACCGCTTAAGTATACCCGCTGGGGA	484
Qy	844	ctatgctgcgaagagtgaactcacaaggaatgaagcggggtccgcaacagcgtctggagca	903

Db 483 GTACGTTGCAAGATGAACTCAAGAAATTGACGGGGCCCGCCACAGCGGTGACGA 424  
 Qy 904 tgggttaataatgatgatatacccaaaactctcgtgggttgacatgagtcgaac- 962  
 Db 423 TGTGTTTAATTCAATGCAAGCGAAGACCTTACAGGGCTTGACATGCCGCAATCT 364  
 Qy 963 atcgagataataagagccttcggg--cagatcaagtgctgcatggtctgcacg 1020  
 Db 363 CTGAAGAAGAGGGGGTGCCTTCGGGAACGCGACACAGGTGTGTCATGCTGCTCAGC 304  
 Qy 1021 tctgtctgagatggttggttaagtcctcgcaagcgcaaccctatcgatgtgct 1080  
 Db 303 TCGTGGCGTAAAGTGTGGGTTAAAGTCCCGACAGAGCGCAACCTCTGTGTTAGTTGCC 244  
 Qy 1081 a-ccttaagtggcactggtgaagaactgcggttgacaaacgagagagagggagtg 1139  
 Db 243 ATCTTGAAGTTTGAACCTGACAGACTGCGGGATGAACCGGAGAGAGTGAAGGATG 184  
 Qy 1140 acgtcaaatccatagccttcatgctcgaaggcaacacagctgtacatgacgataca 1199  
 Db 183 ACGTCAAGTCAATGATGCCCTTATGCCCCGCGACACAGTGTCTACATGCGCGGACA 124  
 Qy 1200 gaaggctgcgaactcgcaagagggagctaatctctaaagtcggtccagttcgatgg 1239  
 Db 123 AAGGATCGCGATCCCGGAGGGGTGAAGTAAACCCCAAAACCCTGCTCATTGGATTGC 64  
 Qy 1260 ggtctgcacacgagcccatgaagtgcgaatcgctagtaacgcgagatcaacg 1311  
 Db 63 AGGCTGCAACTCGCTCGATGACGCCGAATCGTAAATCGCGGTCAGC 12

RESULT 3  
 LOCUS AO957363 716 bp DNA GSS 28-JAN-2000  
 DEFINITION LERAP36TR LERA Arabidopsis thaliana genomic clone LERAP36, DNA sequence.  
 ACCESSION AO957363  
 VERSION AO957363.1 GI:6785064  
 KEYWORDS GSS.  
 SOURCE thale cress.  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eustosids II; Brassicales; Brassicaceae; Arabidopsis.  
 REFERENCE 1 (bases 1 to 716)  
 AUTHORS Buell,C.R., Lin,X., Pai,G., Barnstead,M., Bowman,C., Uterbach,T., Feldblyum,T., Liang,F., Creasy,T., and Fraser,C.M.  
 TITLE Genomic survey sequencing of Landsberg erecta ecotype of Arabidopsis thaliana and identification of sequence-based polymorphisms  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Xiaoying Lin  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: atetlgr.org  
 For additional information, see <http://www.tigr.org/tdb/at/at.html>  
 Seq primer: TR  
 Class: Shotgun.  
 FEATURES  
 source Location/Qualifiers  
 1..716  
 /organism="Arabidopsis thaliana"  
 /strain="Landsberg erecta"  
 /db\_xref="taxon:3702"  
 /clone="LERAP36"  
 /clone\_1id="LERA"  
 /note="Organ: Leaf; Vector: PHOS1; Total genomic DNA was sheared to 0.9-1 Kbp before ligation."  
 BASE COUNT 144 a 213 c 154 g 205 t

Query Match 28.7%: Score 424; DB 13; Length 716;  
 Best Local Similarity 77.6%: Pred. No. 4,36-106;  
 Matches 545; Conservative 0; Mismatches 140; Indels 17; Gaps 2;

Qy 275 gtccgaccacatggaactgagaaacggttcatactcctctacggaaggaagtaagaa 334  
 Db 716 GACCAAGCACATGAGGAAGTAAGACACGGTCCAGACTCTTACGGGAGGACGAGGGGAA 657  
 Qy 335 tctgtctaaatgggggaaacccctgaagcagcagccgctggaaggaagagctctcg 394  
 Db 656 TTTTGGACAATGGGCGCAAGCTGTATCCAGCATTCCTCCGGAGTGAAGAGGCTTCGG 597  
 Qy 395 attgtaagttcatatgagcagaaataaagcaagatgatatgttaactgcta-- 452  
 Db 596 GTTGTAAGCTCTTTCCGAAGGAAAGAACTTACTTTCTAATAAGAGTGAAGCTGACG 537  
 Qy 453 -----aagcaccgctaactacgctgcagcagccgctgtaatacgtatggtg 499  
 Db 536 GTACCTTGATGAAGACACCGGCTAATAGTCCAGCAGCCCGTAAATACGTAGGGTG 477  
 Qy 500 caagcgtgttcggaatcatatggtggttaagagtgctgaagcgaattgttaagtcagtg 559  
 Db 476 CGAGGTTAATCGGAATTAATCTGGGCGTAAAGCGTCCAGCGGCTTGGCAATGATG 417  
 Qy 560 tgaaaactgagcgtcaaccgctgagccttgaactgaactcaagctcgtgaattgagag 619  
 Db 416 TGAATATCCCGACGCTCAACTTGGGAACTGCTTTGAACCTCCACAGTAAATATGTCAG 357  
 Qy 620 aggcagtggaattcctcaggtgtagcgtgaagtcgtagatatactbgaagaaacccagtg 679  
 Db 356 AGGGGGGTGAATTCACAGCTAGACGTAAGATGCGTGAAGATGCGAGATACCAATG 297  
 Qy 680 gcggaagcgaactgctgtgctcaaaactgacgctgagcagaagaagcgtgtagtaacg 739  
 Db 296 CGGAAGCGACGCCCTGTGGATTAATTTGACGCTCTATGACAAAGACGGGGAGAAACA 237  
 Qy 740 ggaatagatccccggtatccacgcgccctaaacgctgtgtcaacgattgtgggggtttta 799  
 Db 236 GGATTAGATACCTCGGTGATGTCACGCCCTAAACGATGTCTAGTGTGTGGTGAATTA 177  
 Qy 800 a--ccctcagtaacgaacctaacggaattagtagacgcgcttgggagactatgctcgca 857  
 Db 176 AATCATATGATGACACACCTAAGCGTGAAGTAGACCCCTGGGGAGTACGCTGCAAGA 117  
 Qy 858 gtgaacccaagaagatgagcgggggttcgcacaaagcgttgagatgtgtttaattcg 917  
 Db 116 TTAATACTCAAGGAATGACGGGGCCCGCAAGCGGTGATTAATGTGATTAATTGCG 57  
 Qy 918 atgatacccaaaaacccaccctggtggttgacatgacatga 959  
 Db 56 ATGCACCGCAAAACCTTACTGCGCTTGACATGCCACTAA 15

RESULT 4  
 LOCUS BF620647 868 bp mRNA EST 22-FEB-2001  
 DEFINITION HVSMEC0020H16f Hordeum vulgare seedling shoot EST library  
 HVCDNA0003 (Etolated and unstressed) Hordeum vulgare cDNA clone  
 HVSMEC0020H16f, mRNA sequence.  
 ACCESSION BF620647  
 VERSION BF620647.2 GI:13109622  
 KEYWORDS EST.  
 SOURCE barley.  
 ORGANISM Hordeum vulgare  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae  
 ; Triticeae; Hordeum.  
 REFERENCE 1 (bases 1 to 868)  
 AUTHORS Wing,R., Close,T.J., Kleinjohs,A., Wise,R., Begum,D., Frisch,D., Yu  
 'Y., Anderson,H., Dale,J., Henry,D., Kennodle,S., Palmer,M., Rambo  
 'T., Saski,C., Schwartzbeck,D., Simmons,J., Choi,D.W., Main,D., and  
 Wood,T.  
 TITLE Development of a genetically and physically anchored EST resource





OY 659 atactctggaggaacccaggctgcggaagcgactgtcgtctcaaaaactaacgttcgaagca 718  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 670 AGATGGAAAGAACACCAACGCGGAAGCACCCTCTGTGTGGCCGACACTGATCGAGA 611  
OY 719 cgaagaatctggatgaataacggatttagatcacccggttaaccacgccctlaacgtttgc 778  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 610 CGAAGCTAGAGGAGCAGTAAGGATTAGTACCAGTAGTGCTTACGCCTTAAACGATGGA 551  
OY 779 taacagtgttgggggtltaaacc--tcagtacaagaacctaaagattaaatagtagcgc 836  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 550 TACTAGGCGCTGTGGTTCGATTCGACCCGTGACAGCTGTACTTAACCGTGTAACTATCCCC 491  
OY 837 ctggggactatgcctcgcgaagtggaacacccaagaatattgaacgggggttcgcgaacagcg 896  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 430 CTGGGAGTACTCTTGCAAGATGAAACTCAAAGGATTGACGGGGGGCCCGACAACGG 431  
OY 897 tggagcatgttgtttaactcgaatgaacccccaaaaaacctcaccttggcttgatcattgatc 956  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 430 TGGACCATGTGGTTTAATTCGATGCAGAAACGAAACCTTACCAGGGGCTTGACATGCCG 371  
OY 957 tgaatc-atgtagaatatatagaaccttcgggcagat--tcacaggtctcatgtgtt 1013  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 370 GAATCCTCTTGAAGAAAGAGGGGTGCTTCGGGAACGCGGACACAGTGGTGCATGGCTGT 311  
OY 1014 cgtcagctcgtatctgtgaatgtttggttaatgccgaacagcgcaacccctatcgt 1073  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 310 CCTCAGCTCTGTCCTTAAGCTTTGGGTTAATCCCGCACACGCCCACCTCGTGT 251  
OY 1074 tgttgcta-ccctaaagttgggacatcgtgaacgaactcgcggtlgacaaaacggagaagcg 1132  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 250 AGTTCCCATCTGTTAGTTTGAACCCCTGMAACAGACTGCCGGTGATNAGCCGAGGAAGGT 191  
OY 1133 gggagtgacgttaaatctctcatatgaccttatgtlccaagggccacacagtgctacaatggc 1192  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 190 GAGGATGACGCTCAATCATCATGCCCCCTTATCCCTGGGCGACACAGCTCATCAATGGC 131  
OY 1193 gcaatacgaaggtgcgcgaactcgaagaagagatcaatctcfaaagttgcgtccagttc 1252  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 130 CCGGACAAAAGGTCCTCGATCCCGCGAGGCTGAGCTNACCCCAAAAACCTGCTCTCACTTC 71  
OY 1253 ggattgggtcttgcacactcgaacccatgaagtgcggaatcgtctagtaactcgcgcatcagc 1311  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 70 GGATTCAGAGCTGCAACTGCTCGATGATAAGCCGGAATGCTAGTAAATGCGCGGTACG 12

RESULT 6  
CNS00MCV 720 bp DNA GSS 28-JUN-1999  
LOCUS CNS00MCV Arabidopsis thaliana genome survey sequence SP6 end of BAC T1H14  
DEFINITION OF T1M1 library from strain Columbia of Arabidopsis thaliana,  
genomic survey sequence.  
ACCESSION AL093421  
VERSION AL093421.1 GI:5294575  
KEYWORDS GSS.  
SOURCE thale cress.  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
REFERENCE 1 (bases 1..to 720)  
Salanoubat,M., Choisne,N., Artiguenave,F., Brottier,P., Wincker,P.,  
Samson,D., Saurin,W., Weissenbach,J. and Quetler,F.  
JOURNL Unpublished  
REFERENCE 2 (bases 1 to 720)  
GENOSCOPE.  
AUTHORS Direct Submission  
TITLE Submitted (25-JUN-1999) Genoscope - Centre National de Sequencage :  
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
FEATURES  
Source Location/Qualifiers  
1..720 /organism="Arabidopsis thaliana"

Query Match	27.6%	Score 407.4	DB 13	Length 720
Best Local Similarity	77.2%	Pred. No. 1.6e-101		
Matches 547	Conservative 0	Mismatches 156	Indels 6	Gaps 4
QY 609	ggagtttggaagagcgaatggaattccaggtgtcagcgtgtgaatgcgttaatatcttgag	668		
Db 1	GAGTACGGTTGGGGCGAGGGAATTTCCGGTGGAGCGGTAAATGCGTAAAGATCGGAAA	60		
QY 669	gaacaccaatgagcgaaggaaggaactgtctgtgtcgaacatgacgcgtgaagcagaagcgtg	728		
Db 61	GAAACACCAACGGCGAAGAGCACTGCTGGGCCACACTGACTGAGAGACGAAGAAGTAG	120		
QY 729	ggtagtaaacgagatagatatacccggtaaatccacgccttaaacgtgtctacaaagttgt	788		
Db 121	GGAGCGGAATGGGATTAAGATTACCCCGTAGTACTAGCCGTAAGAGATGATACGCGC	180		
QY 789	tgggggttttaaccc--tcagtaacgaacttaacggaattagtagaccgcttggagacta	846		
Db 181	TGTCCTATCGACCCCTGCGAGTCTGTAGCTAAGCGCTTAAGATCCCGCTGGGAGTA	240		
QY 847	tgctcgcaagatggaacccaagaagaattgaaggggtccgcacaagcgttgagacatgt	906		
Db 241	CGTTCCGAAGAAGAAATCTCAAGAAATTACGGGGGCCCGCACACAGCGGTGACGATCT	300		
QY 907	ggtttaattcgatgatalaccaccaaaactccactcgtgggtctgacatggaatcgaatcagt	966		
Db 301	GGTTTATTTGATGCAAGCGAAGAAACCTTACACGAGGCTTGACATGCGCGCAATCCTCTT	360		
QY 967	-agagatatagagccttcgggcagat--tcacagtgctgcatagtgtgtcgtcaagctcg	1023		
Db 361	GAAAGAGAGGGGGCTTCGGGGAACCGGACACAGGTGTGTGCTGTCTGTCACGCTCG	420		
QY 1024	tgctcggaagatgttggttaagtcctcgcaagagcgcaaccctatcgtatgtgtgtcagc	1083		
Db 421	TGCGCGAAGGTGTGGGTTTAGTCCGCAACGAGCCGACACCTCTGCTTTAGTTGCCACC	480		
QY 1084	-ttaaattggaactgtgtcgaacatgcgcgtgtgacaaaccggaaggaagcgcggaatagc	1142		
Db 481	GTTGATTTGGAAACCTTGAAACAGACTGCGCGTATTAAGCGGAGAAAGGTGAGGATACG	540		
QY 1143	tcaaatctcaatgagccttattgtccagggccacacacgltcctaattgctcgatatacag	1202		
Db 541	TCAAGTCATCATGCCCCCTTATGCCCTGGGGGACACAGTCTACAAAGGCGCGGACAAAG	600		
QY 1203	ggttcgcaactcgcgaagagaggaactaattccttaaaagtctgtccagttcgtatgtggt	1262		
Db 601	GGTCGGATCCCGCGAGGGGAGCTTAATCTCAAAAAACCCCTTCACTTGAGATTGCGAG	660		
QY 1263	ctgcaactcgaaccacaatgaagtgcgaatcgtatgaatcgcggaatcagc	1311		
Db 661	CTGCAACTCGCGCTGCATGAAGCGGAAATGCTTGAATGCCCGGTACGC	709		
RESULT 7	BH016730	755 bp	DNA	GSS
LOCUS	BH016730/c			18-MAY-2001
DEFINITION	TDGCP25TH CTG Lycopersicon esculentum genomic clone CTOG17F2, DNA			
ACCESSION	BH016730			
VERSION	BH016730.1	GI:14147768		
KEYWORDS	GSS.			
SOURCE	tomato.			
ORGANISM	Lycopersicon esculentum			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			



Oy	204	ccggcgacgagatgaactgtcttgcttgcaggtaaaatgacctcaagaagcagcagtcggttaaccgg	263
Db	119	CGAATCTGATTACTAGTTGGTGGGTAAAGCCTAACCAAGGCTTCGACATCACTACTTG	178
Oy	264	ccttagagggatgtccgcgcacaatgtaaacctgaagacagcgtlccalactctlaagygagca	323
Db	179	TCTGAGAGGAGCAGACCACCACTACGTGAACCTGAGCACAGGCTCCAGACTCCTCAGCGGAGCA	238
Oy	324	gcagtttaagaattcttgtccaatgaggggaaaacctgaagcagcgaagcgcgcgtgaacgaag	383
Db	239	GCACTGGGGGAATTTTTGGAACAATGGGCCCAAGCCCTGATCCAGCATTCGCCGTGAGGAAG	298
Oy	384	aagcttcctgatgttaaagtccaatgagcaggaaaaataagcagcaatgtgaatgtgt	443
Db	299	AAGGCCTTCGGGTTGTAAACCTCTTTCCGACAGGAGGAAAACCTTACTTCTTATAAAGAG	358
Oy	444	acctgtcta-----aagcacccgctaactacgtgtccagcagcccggttaa	488
Db	359	TGAGCGTGCACGCTACCTTGTATGAAACACACCGGCTTAATCTGTCACAGACGCGCGGTAA	418
Oy	489	tagtatgtgtcaagcgtltgttcgaaatcatgtggcgttaaggtgtcgtgagcgagattg	548
Db	419	TACCTAGGCTCCGAGCCTTAATCGGAATTACTGGCGCTAAAGCGTGCAGCGGCTTTGG	478
Oy	549	taagtcaggtgtgaanaactcgcgcgtcaaccgcctgagccttgcacttgaactaacaatctg	608
Db	479	CMACTCAGATGTGAAAATCCCCGAGACTCAATTGGGAACCTCGTTGAACTGCGACACTA	538
Oy	609	gagtttggagagagcgaagtgtgaattccaaggtgtatgagcgtgaatatgtgtatctcgag	668
Db	539	GAATATGTCAAGAGGGGGGTAGAATTTCCACTGTGACGTAATAATGCTAAGATGGTAGAG-598	
Oy	669	gaacacacgtgtgcgaagcgacactgtctgcgtccaacactgaacgtctgagcagcaaaacgtg	728
Db	599	GAATACCAATGCGCAAGSCAGACCCCCTGNGATATATATGACGCTCATGACAGAAACGTG	658
Oy	729	ggtgtgtaaacgggatttgataccccggttaattccaagcctcaaatcgtltgtctacagttgt	788
Db	659	GGGAGCAAACGAGTATGATACCTGTAATCCATCCACGCCCTTAACAGATGCTACTAAGTTGT	718
Oy	789	tgggggttttaa	800
Db	719	TGCTGAGAGTAA	730
RESULT	9		
AO957181/c			
LOCUS	741 bp	DNA	SSS
DEFINITION	LERRAO3JTR LERA Arabidopsis thaliana genomic clone LERRAO3, DNA		
ACCESSION	AO957181		
VERSION	AO957181.1	GI:6784882	
KEYWORDS	GSS.		
SOURCE	thale cress.		
ORGANISM	Arabidopsis thaliana		
REFERENCE	Eukaryotes: Viridiplantae; Streptophyta: Embryophyta: Tracheophyta		
AUTHORS	Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots;		
TITLE	Rosidae; eurosoids II; Brassicales; Brassicaceae; Arabidopsis.		
JOURNAL	Buell,C.R., Lin,X., Pal,G., Barnstead,M., Bowman,C., Uteirbach,T.		
COMMENT	Feldjbaum,T., Liang,F., Cressy,T. and Fraser,C.M.		
	Genomic survey sequencing of Landsberg erecta ecotype of		
	Arabidopsis thaliana and identification of sequence-based		
	polymorphisms		
	Unpublished (2000)		
	Contact: Xiaoying Lin		
	The Institute for Genomic Research		
	9712 Medical Center Dr., Rockville, MD 20850, USA		
	Tel.: 301 838 0200		
	Fax: 301 838 0208		
	Email: atetig@org		
	For additional information, see http://www.tigr.org/tdb/at.at.html		
	Similar to A. thaliana chloroplast sequence (GB:AF000423)		

[illegible]

Accession	Version	Keywords	Source	Organism	Reference	Authors	Title	Journal	Comment
AA0957182	1	GI:6784883	GSS	Arabidopsis thaliana	Buell, C.R., Lin, X., Pai, G., Barnstead, M., Bowman, C., Uterbach, T., Feldblyum, T., Liang, F., Cressy, T., and Fraser, C.M.	Genomic survey sequencing of Landsberg erecta ecotype of Arabidopsis thaliana and identification of sequence-based polymorphisms	Unpublished (2000)	Contact: Xiaoying Lin The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 0200 Fax: 301 838 0208 Email: acet@iglr.org For additional information, see <a href="http://www.iglr.org/tcd/at/at.html">http://www.iglr.org/tcd/at/at.html</a> Similar to A. thaliana chloroplast sequence (CB:AP000423) Seq primer: TR Class: Shocgun.	
FEATURES	source	Location/Qualifiers							
1..757	/organism="Arabidopsis thaliana"								
/db_xref="taxon:3702"									
/clone="LERA034"									
/note="Organ: Leaf; Vector: pHOS1. Total genomic DNA was sheared to 0.9-1 kbp before ligation."									
BASE COUNT	137 a	247 c	170 g	203 t					
ORIGIN									
Query Match	25.6%	Score 378.8	DB 13	Length 757					
Best Local Similarity	74.6%	Pred. No. 1.2e-93							
Matches 519; Conservative	0	Mismatches 167	Indels 10	Gaps 3					
190	gcttgtagatgtagcccgccgagccattagctatgtgtgtagtaatgctcaccaagcgga	249							
689	gccccagagagggcctgctgctcttattagctagtgtgtgagccaatgcttaccagagca	630							
250	cgatcgtagtcgagccgtgtagaggtgtccgagccaaatggaatgtagaacggtctcatc	309							
629	tgatcagtagtgcgtgctccgagcattatcagccacacctggagtagacagccccaagac	570							
310	tcttaaggagagagagcagatlaagaatcttgcctcaatcgggggaaacctgaagcagcga	369							
569	tctttagggagagcagacagctgagagagagagagagagagagagagagagagagagag	510							
370	ccgctgtagaagaaagagcttcgcgaltgttaagttcaatgaagcaggaataaataagcagc	429							
509	ccgctgtagaag	452							
430	aagtgtagtgtgttactgcttaagagacccggttaactgtagtgcagcagccgctgaat	489							
451	gagcgtatctgggga-----atnagcattggccttaactctgtgtccacagcccggtaat	398							
490	acgtatgtagcagaagcgtgttcggaatcatctggtggtggtggtggtggtggtggtggt	549							
397	acagagagatgcaacgcttattccggaatgattggcgctgaaagcgtctgtagtgcttttt	338							
550	aagtcaggtgtgaactgtagcgtgcttaacccgtgtgctgtcaactggaactacacgaagct	609							
337	aagtcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag	278							
610	agtttgtag	669							
277	agtagcag	218							

QY	670	aacacacgctgaggaagcggaactgctgctgctccaaactgacgctgagcgacggaagcg	729
Db	217	AACCAACACGCGCCAAAGCAGCTGCTGGGCGCCAGCACTGAGACTGAAGACGAAAGCTAGG	158
QY	730	gtatgaacgaggaatataaccgcytaatccacgcgcctaaacgctgtctaccagttgt	789
Db	157	GGAGCGAATGGAGTATACATACCCAGTAGCTTAGCCGTTAAACGATGAGTACTAGCGCT	98
QY	790	gaggggtttaacc--tcagtaacgaacctaaagatgaagtgaaccgctggagacct	847
Db	97	GTGGGTATCGACCGCTCACTGCTGATGCTAACGCGCTTAAGTATCCCGCTGGGGAGTAC	38
QY	848	gctgcgaagatgaactcaaaagaaatgacgaggg	883
Db	37	GTTCGCAAGATGAAGTCAAGGAAATTTGACGGGG	2
RESULT	11		
LOCUS	B73824	676 bp	DNA
DEFINITION	T27K4/R. TAMU Arabidopsis thaliana genomic clone T27K4, DNA		
ACCESSION	B73824		
VERSION	B73824.1		
KEYWORDS	GI:2712975		
SOURCE	GSS.		
ORGANISM	thale cress. Arabidopsis thaliana Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidops. 1 (bases 1 to 676) Rounsley,S.D., Field,C.E., Bass,S., Linher,K., Linher,K., Golden,K. , Berry,K., Granger,D., Suh,E., Wible,C., Adams,M.D. and Venter ,J.C. A BAC End Sequence Database for Identifying Minimal Overlaps in Arabidopsis Genomic Sequencing. Update 3 Unpublished (1997) Other_GSSs: T27K4TF		
JOURNAL	CONTACT: Steve Rounsley Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 0200 Fax: 301 838 0208 Email: rounsley@ligr.org Seq primer: M13 Reverse Class: BAC ends		
COMMENT	High quality sequence stop: 676. Location/Qualifiers 1..676 /organism="Arabidopsis thaliana" /strain="Columbia" /db_xref="taxon:3702" /clone="T27K4" /clone_1lb="TAMU" /sex="hermaphrodite" /note="vector: BelobacII; site_1: HindIII; site_2: HindIII ; Produced by Rod Wing"		
FEATURES			
source			
BASE COUNT	170 a 164 c 213 g 128 t	1 others	
ORIGIN			
Query Match	25.5%	Score 376.6; DB 13; Length 676;	
Best Local Similarity	76.2%	Pred No. 4.8e-93;	
Matches	515; Conservative 0; Mismatches 155; Indels 6; Gaps 4;		
QY	610	agtttggaagaaagaaatccacgtgtgacggtgaaatcgctagataltcggag	669
Db	1	AGTACGCTAGGCGGAGGAGGAATTTCCGCTGGGAGCGGTAAATGCTAGAGATCGGAAAG	60
QY	670	aacacacgctgaggaagcggaactgctgctccaaactgacgctgagcgacggaagcg	729

```

Db 61 AACACCAAGCGGAGAACACTCTGCTGGCCGACACTGACACTGAGACGAAAGCTAG 120
Oy 730 gtaataaacgggattagataaccggtgtaataccacgcttaaacgtgtctaccagttgt 789
Db 121 GGAGCGAATGGGATTAGTACCCAGTAGTCCCTTACCCCTAAACGATGATACTAGGCCCT 180
Oy 790 ggggggtttaaacc--taagtaacgaacctaaacggattaaagtaacgccttggggacta 847
Db 181 GTGGCTATCGAACCCGCTGACAGTCTGCTACTAACCGGCTTAAGATACCCGCTGGGAGTAC 240
Oy 848 gctgcgaagaatgaaactcaagaagaattgacgggggttccgcgaacaggcggttgagcaatg 907
Db 241 GTTGCGCAAGAAATGAAACCAAGAAATTAACGCGGCGCCGCAACAGCGGTGAGCATGTG 300
Oy 908 gtttaattcgaatgatacccaaaaacctcacctggggttgaatgataatc-atgtc 966
Db 301 GTTAAATTCGATGCAAAACGAAAGAACTTACAGGGCTTGAATGCGCGGAATCTCTTGG 360
Oy 967 agagataatgagcccttcggg--cagattcaaggctgctgaatgtgtcgtcaagctgt 1024
Db 361 AAAGAGAGGGGGCTTGGGAGACGCGACACAGGTGGTGCATGCTGCTGACACTGCTGT 420
Oy 1025 gtcgtgagatgttgggttaagtcgccgaacgagcgcaacccctatcgtatgttgaacc- 1083
Db 421 GCCGTAAAGCTGTGGGTTAAGTCCCGCAACGAGCCCAACCTCGTGTAGTCCACCG 480
Oy 1084 ttaagtgggacacgtgacgaacctgcgggtgacaaacggagagagagcggtgagact 1143
Db 481 TTGAGTTTGGAAACCTTGAAACAGACTGCCGATGATAGCCGGAGGAGTGAGACATGAGCT 540
Oy 1144 caaatccatcagcccttcatgtccagggccacacagctgtcaatctgagcgaatacagag 1203
Db 541 CAAGTCATCATGACCCCTTATATCCCTGNGCAGACACAGTGTACATATGGCCGGGACAAAG 600
Oy 1204 gtccgcaactcgcaagagagagtaactctcaaaagtcgtgccagttcgatgtgggtc 1263
Db 601 GTCCGGAATCCCGGAGGGTGAGCTTAACCTCAAAACCCGCTCTAGTTGCGATTCAGAGC 660
Oy 1264 tgcgaactcgaccocat 1279
Db 661 TGCAACTCGCTGCAT 676

RESULT 12
BH016557/c 655 bp DNA GSS 18-MAY-2001
LOCUS BH016557
DEFINITION TDCN35TH ctog Lycopersicon esculentum genomic clone ctog17E22, DNA
sequence.
ACCESSION BH016557
VERSION BH016557.1 GI:14147595
KEYWORDS GSS.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asterales; euasterids I; Solanales; Solanales; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 655)
AUTHORS van der Hoeven,R., Sun,H., Cho,J., Utterback,T., Ronning,C. and
Tanksley,S.
Tomato Demethylated Genomic DNA sequences
Unpublished (2001)
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
tomato demethylated genomic DNA
Seq primer: M13F-R
Class: shotgun.
FEATURES
Location/Qualifiers
1..655
/organism="Lycopersicon esculentum"

```

```

/cultivar="E6203"
/db_xref="taxon:4081"
/clone="ctog17E22"
/clone_lib="ctog"
/rissue_type="young leaves"
/dev_stage="12-14 weeks post harvest"
/lab_host="E.coli JM109"
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; This library was made from short EcoRI digested
fragments of the genome of Lycopersicon esculentum ligated
into pBS (SK-). The fragments were cloned into the
methylation restrictive E.coli strain JM109 with the
purpose of enriching the library for non-methylated DNA
fragments. This procedure may enrich the pool of cloned
fragments in JM109 cells for sequences representing
expressed genes. Average insert size 1.27 kb."

BASE COUNT 125 a 203 c 168 g 159 t
ORIGIN

Query Match 24.7% Score 364.8; DB 13; Length 655;
Best Local Similarity 77.0%; Pred. No. 8.4e-90;
Matches 496; Conservative 0; Mismatches 142; Indels 6; Gaps 4;

Oy 674 ccagtgcggaagcgactgtgctgctcaaaactgagcgttgagcgcaagaagctggtag 733
Db 655 CCAACGGCGGAAGCACTCTGCTGGGCGCACACTGACAGAGCAAGAACTAGGGAG 596
Oy 734 taacggaattagatacccggttaataccagccctaaacgtgtctaccagttgtggg 793
Db 595 CGAATGGGATTGATTAACCCAGTACTTACCGGTAACAGATGATGATAGCGCGCTGTC 536
Oy 794 gtttaaccc--tcagtaacgaactaaacgattagatagacgcgttgggactatgtc 851
Db 535 GTATGACCCGTCGTCGTAACGCGTAAGATGATGATGATGATGATGATGATGATGATG 476
Oy 852 gcaagatgaaactcaagaagaaatgaacgggggttcgcgaacgggtgagcatgtgtt 911
Db 475 GCAGAAATGAATCAAAAGAAATTAACGCGGCGCCGCGCAACGCGTGAAGATGGGTTT 416
Oy 912 aattcgaatgaatacccaaaaacccacactgactgtgacatgataatc-atgtagag 970
Db 415 AATTGATGCAAAAGCAAGAAACCTTACCGGCTTGAATGCGCGGAATCTCTTGAAG 356
Oy 971 atatatgagcccttggg--cagattcaaggtgacgtgtgtgtgtgtgtgtgtgtgtgt 1028
Db 355 AGAGGGGTGCTTCGGGAAACGCGGACACAGGTGTCATGCTGTCATGCTGTCGCG 296
Oy 1029 tgaatgtgtgttaagtcgccgaacgagcgcaacccctatcgtatgtgtcta-ccctaa 1087
Db 295 TAAGGTGTTGGGTTAAGTCCGCAACGAGCGCAACCCCTGCTTAAATGTCATGCTTGA 236
Oy 1088 gtgggcaactgtgtacgaaactgtcggtgacaaaacggagaaagcggtgataagctcaa 1147
Db 235 GTTTCGAACCCCTGAACACACTGCGCGGTATTAAGCCGGAAGGATGAGATGATGCTCAAG 176
Oy 1148 tccatagcccttattgcccaggggccacacacgtgtctcaatagtcgagataagaggtgtc 1207
Db 175 TCATCATGCCCCCTTATGCCCCGCGACACAGTCTCAATAGGCGCGGACAAAGGCTG 116
Oy 1208 ccaactcgcaagagaggaactaactctaaagtcggtcccaagttcgatgtgggtctgtca 1267
Db 115 CGATCCGCGAGGAGGTGAGCTTAACCCCAAAACCCGTCCTCAATTCGGATTCGAGCTGCA 56
Oy 1268 actgagcccatgaaagtgcggaatcgctagtaatcgcgatcagc 1311
Db 55 ACTCGCTGCATGAAGACGCGAATGCTGTGATATCCCGGTGAGC 12

RESULT 13
AO956298 693 bp DNA GSS 28-JAN-2000
LOCUS AO956298/c
DEFINITION LERAJ19TR LERA Arabidopsis thaliana genomic clone LERAJ19, DNA

```





Mon Dec 17 07:48:13 2001

us-09-380-826a-1.rst

Page 12